Yale Bicentennial Publications

LIGHT

Yale Bicentennial Publications

With the approval of the President and Fellows of Yale University, a series of volumes has been prepared by a number of the Professors and Instructors, to be issued in connection with the Bicentennial Anniversary, as a partial indication of the character of the studies in which the University teachers are engaged.

This series of volumes is respectfully dedicated to

The Graduates of the University

LIGHT

A CONSIDERATION OF THE MORE FAMILIAR PHENOMENA OF OPTICS

BY

CHARLES S. HASTINGS, PH. D.

Professor of Physics in Yale University

NEW YORK: CHARLES SCRIBNER'S SONS

LONDON: EDWARD ARNOLD

1901

Published, August, 1901

UNIVERSITY PRESS · JOHN WILSON
AND SON · CAMBRIDGE, U.S.A.

ca.

PREFACE

There is a very large number of phenomena in the domain of optics which may be regarded as familiar, in the sense that all of them may be observed without the accessories which are found only in the collections of philosophical apparatus characteristic of physical laboratories. To describe and explain this class of phenomena is the aim of this book. The extensive subjects of spectroscopy and of polarized light are excluded by the limitation imposed; the remaining topics are presented rather with a view to usefulness than with respect to any arbitrary notion of relative importance. For the reason just given the phenomena of color sensations are dealt with in a rather brief way, since that admirable work, the "Students' Text-Book of Color," by Professor Rood, supplies, far better than a less accomplished writer could hope to do, all demands of the ordinary reader. On the other hand, so little has been contributed during more than half a century to what may be called atmospheric optics, that it seems justifiable to extend the chapter treating this subject to an unusual degree.

Very great improvements in the theory and construction of the most important optical instruments have been made since any popular work devoted to their consideration has appeared. Indeed, no branch of applied science has experienced a more remarkable revolution in recent times than that which the advance of theoretical optics has impressed upon the skilled

optician of the present. The history and meaning of this change is given in the chapters on the telescope and on the microscope.

Those readers who are fortunate enough to command the effective aid of mathematical language will find in the appendices many rigid and easy proofs of important optical theorems which must, in the text, necessarily rest upon more diffuse reasoning. It is hoped that the extraordinary power of the method of analysis presented in Appendix A will recommend it to the student of physics.

CONTENTS

CHAPTER I

WAVE MOTION — REFLECTION — REFRACTION

CHAPTER II

OPTICAL INSTRUMENTS

CHAPTER III

PHENOMENA OF LIMITED WAVE-SURFACES — INTERFERENCE — WAVELENGTHS OF LIGHT

CHAPTER IV

DISPERSION — CHROMATIC EFFECTS OF DIFFERING WAVELENGTHS — COLORS OF THIN PLATES

CHAPTER V

THE TELESCOPE

CHAPTER VI

THE MICROSCOPE

CHAPTER VII

OPTICAL PHENOMENA OF THE ATMOSPHERE

CHAPTER VIII

THE EYE AND VISION

CHAPTER IX

APPENDIX A

APPENDIX B

APPENDIX C

LIGHT

CHAPTER I

WAVE MOTION — REFLECTION — REFRACTION

THE word *light* is used in two distinct senses, namely, to designate the sensation which is characteristic of the organ of vision, and also as a name for the usual cause of that sensation. This double meaning of the word would result in little inconvenience if there were always a definite relation between the sensation and its cause; but this is far from being true. For example, when we speak of white light, we may mean a certain sensation which is perfectly definite and familiar to all seeing persons, or we may mean that form of energy which can give rise to such a sensation. In this second sense the term is wholly indefinite, since there is an infinite variety of forms of energy which may give rise to the sensation of whiteness. The difficulty, which is a serious one in scientific language, may be avoided by restricting the meaning of the word to one of its significations, preferably to that of a sensation, after the analogy of the use of the word sound. But such a restriction would not be in accordance with well-established usage, and would necessitate the frequent employment of awkward circumlocutions. Another means of avoiding confusion is so to divide the subjects treated that the sense in which the word is used is unmistakable. This method has the advantage of conciseness as well as that of being in accordance with the usage of most writers. We shall, then, divide this book into parts, in the first of which light will be treated as a phenomenon of wave motion wholly independent of the sense organ which betrays

its existence to us; in the second part we shall regard light as a sensation, and consider more particularly the relation of the wave motion to those sensations to which we attach definite names. In the first part the eye is simply an optical instrument quite like the photographic camera; in the second part we must study the construction and office of the retina, and the relation of the mental impressions to the character of its stimulations.

We are now able to define perfectly the terms "white light," "yellow light," etc. In the first place, by white light we mean such waves as are emitted by a solid body at a very high temperature, as, for example, the incandescent lime in the lime-light. Any other kinds of waves, even if indistinguishable from these by the unassisted eye, are not white light. Again, yellow light, green light, etc., are the *simplest* waves which will excite in a normal retina the sensations yellow, green, etc. In the second place, on the other hand, these color names designate certain familiar sensations which, as we shall see, may be produced in an indefinite variety of ways.

It is now a little more than two centuries since the Dutch philosopher Huyghens published a paper in which he explained the familiar phenomena of light by waves in a medium that pervades all space and is called the luminiferous ether. His reasoning was so convincing, the explanations so simple, and the experiments supporting his views so apt, that except for the labors of the single philosopher then living, who was greater than Huyghens himself, they could hardly have failed to receive at an early day the universal acceptance which they now command. Nine years earlier, in 1669, Newton had commenced his labors in the field of optics, by which, largely on account of fame and authority won in the domain of mechanics and astronomy, he established a theory of light which remained almost unquestioned for nearly a century and a half. Newton supposed light to consist in extremely small particles of matter projected from shining bodies with enormous velocities. We now know

that this hypothesis was not only less fruitful than that of Huyghens, but, even with the comparatively limited range of optical phenomena known to Newton and his contemporaries, was also less probable. It was left to Fresnel — in a remarkable series of papers of the highest order of merit, extending from 1815 to 1826 — to establish the wave theory upon a foundation which leaves no room for doubt. It is true that there are many questions remaining, both as to the character of the medium by which the waves are transmitted, and the relations which the medium bears to forms of matter more familiar to our senses; but there is no more likelihood of physical science in the future rejecting the essential features of the wave theory of light than there is of the rejection of the doctrine of universal gravitation, or any other of the established laws of mechanics.

As the aim of this book is to explain the more familiar phenomena of light which do not depend on the *form* of light waves, but only on their lengths and velocities, we may acquire all the knowledge of wave motion necessary for our purposes by a study of the changes in the surface of a liquid, such as water or mercury, when agitated by a system or by systems of waves.

If a pebble is thrown into a still pool, it will give rise to a series of circular waves having their centre at the place of original disturbance, and each increasing in diameter at an unchanging rate. The velocity with which the waves move outward from the centre is called the wave velocity; a distance equal to that from one crest to the next is called the wavelength; and the time in which a wave advances its own length is called the period of the wave. These terms are applicable to all types of waves, whether those in air which produce the sensation of sound, or those in the light ether which give rise to the special sensation of the organ of vision. Only one other term is necessary to define all the properties of a system of waves required for our purposes. Half the height of the waves, that is, half the difference of level between the tops of the waves and the bottoms of the

troughs, is manifestly the greatest distance that a particle of water in the surface departs from its original position; this is called the amplitude of the wave. It is obvious that the amplitude of the circular waves under consideration continually diminishes, until at a very great distance it becomes insensible.

If two pebbles are thrown into the pool at the same instant, each becomes the centre of a system of waves which moves outward from its centre exactly as though it alone existed. This peculiarity of independent existence is the most characteristic feature of waves, and is absolutely without limitation. Thus, on the surface of the ocean we may have, and in general do have, a vast number of distinct and independent systems of waves; there are the three systems of tidal waves — the semi-monthly, diurnal, and semi-diurnal; the waves produced by the local winds; those which have had a similar origin at a great distance and pursue a different direction of motion; waves reflected and generated by a passing ship, and every ripple caused by bird or fish as perfectly preserved as on still water. We must not conclude, however, that after our two pebbles are thrown every point of the surface is disturbed twice as much as it would be with a single centre of disturbance, although the whole effect would be obviously twice as great. There would be places where the crests of one system of waves would arrive at the same moment as the troughs of the other system, and if the two systems were at that point of equal amplitudes, the crest of the one would just suffice to fill the trough of the other, thus leaving in a limited area a flat surface without waves. In other places, again, one system would act in conjunction with the other and produce waves of twice the height. This second characteristic feature of wave motion is called *interference*. It can be readily observed in all types of waves. As a common example, we may detect regions of silence about a sounding tuning-fork if it is held in close proximity to the ear and there rotated about its axis. It is true that here the phenomenon is somewhat more complicated; but we

may say, in effect, that the system of waves from one tine of the fork interferes with that from the other, and hence, as regards the air, the two tines bear much the same relation that the two pebbles do in respect to the surface of the pool. So, on the other hand, if, about two centres of emanations of any kind, we find regions where the effects are mutually destructive, interspersed with others where they are additive, we may be sure that such emanations are waves. Just this kind of evidence demonstrates to us that light is a wave motion, as will be shown further on in our study.

If we inquire why the point where the pebble entered the water should be a centre of a whole system of waves, instead of a single one only, we may find the answer in a consideration of this kind: The surface of the water is momentarily depressed where the stone enters it, but immediately the stone passes through, and ceases to act upon the surface, which quickly returns to its original form by the action of weight. When, however, it reaches this original form, it possesses considerable velocity as a result of the continuous action of weight, and this velocity carries the surface beyond its primary condition of flatness, producing a momentary elevation, which, in turn, must fall and send out another wave. But this consideration logically forces us to the conclusion that every point of the water surface which is disturbed must become the centre of its own system of waves, that is, that every point of every wave must be regarded as such a centre.

It was the recognition of this last fact and its pursuit to its legitimate conclusions which constituted the notable discovery of Huyghens, known in the science of optics as Huyghens's principle. This principle is of such importance, and will be so useful to us in our further studies, that it may be profitably discussed here with some care.

Let *C*, Figure 1, be a centre of disturbances, and *ab* a wave from it; what will be the final effect at any other point, as *P*? According to the principle of Huyghens, every point of *ab* must be regarded as a centre of a wave which will finally reach *P*. From *P* draw lines to different points

of ab, as Po_1, Po_2, Po_3, etc., such that each succeeding line is just one-half wavelength longer. It is obvious that the system from the point o_4 will be one-half wavelength behind the system from o_3 when they reach the point P, the crests of the first system corresponding with the troughs of the second; and thus, according to the law of interference, their effects at P will be mutually destructive. So also for the waves from o_2 and o_1. Further, it is obvious that the effect of each point of the wave between o_4 and o_3 will be completely destroyed by the effect of a certain point between o_3 and o_2,

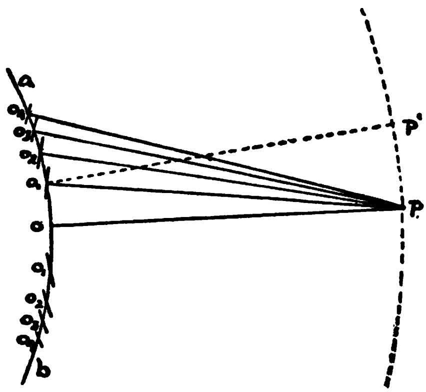

FIGURE 1.

and so on. Since, however, as we approach o the length of the segments into which the wave is divided by the lines from P increases, there is an unbalanced effect of the central segments, and the resulting effect at P is quite the same as though the wave moved on parallel to itself, and only the portion at o were considered. It is very important to observe that if the length of the wave ab were not large compared with the wavelength, or if the wave were interrupted so that disturbances from all points of ab could not reach P, the reasoning and its conclusions would fall to the ground.

We see from this consideration why it is that in a medium which does not change its character the wave moves on

parallel to itself, the disturbance at *o* being propagated along the line *Co* to *P*, and that at o_1 along the line Co_1 to P', etc. If the waves are light waves, these straight lines of propagation are called ***rays*** of light.

After having established this fertile principle, Huyghens could readily explain the phenomenon of reflection as follows:—

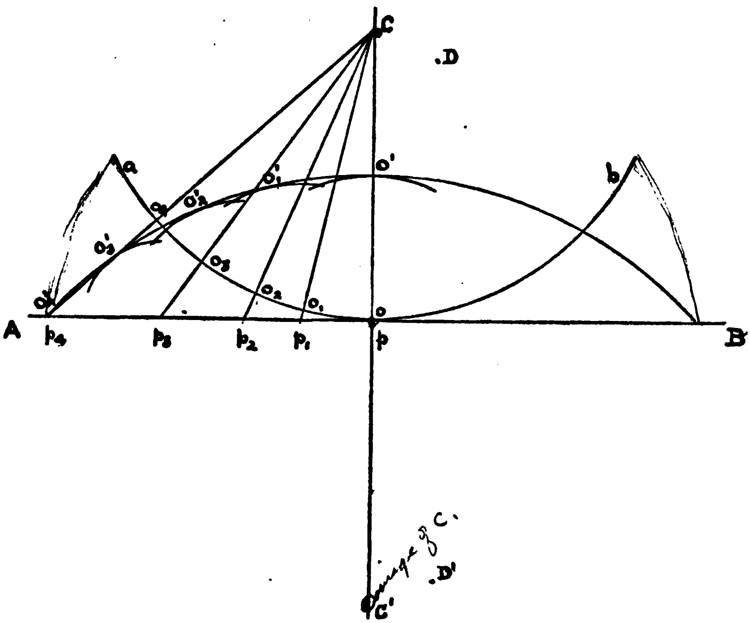

FIGURE 2.

Let *C* in Figure 2 be a centre of wave motion and *AB* represent a mirror, that is, a barrier to the further progress of the waves which does not destroy their motion. Let us consider the state of things at the moment when a certain wave *ab* first touches the mirror at a point *o*. The wave can no longer move on in its original direction; but as, in accordance with Huyghens's principle, the barrier does not destroy the motion, *o* must be regarded as a centre of

disturbances which are propagated in all directions except through the mirror. At a certain instant afterward, the disturbance from o will be found at o' in the circle whose centre is at p. A moment after the wave touches the mirror at o the point o_1 in the wave will reach the mirror at p_1, which in turn becomes a new centre of a circular wave. This disturbance will be found in the circle o'_1 at the instant when the wave from p is in o', since $p_1 o'_1$ is taken less than po' by the distance p_1o_1. In a similar manner we may construct the waves from any number of points on the mirror. In the figure, five such waves are given, the last, o'_4, corresponding with p_4. Hence the reflected wave will be found at the given instant to be a circular wave $o'_4\ o'_2\ o'$, having C' as its centre, as is quite apparent from the construction. This point C', which is the centre of the reflected waves, is called the *image* of C. If by any means the wave is so modified as really to have a second point of divergence, that point is called a *real image;* but if, as in this case, it only seems to come from such a point, it is called a *virtual image.*

This construction is perfectly general, subject only to the condition that the mirror shall be large compared to the wavelength; and it is important not only because it gives a complete definition of an optical image, but also because it contains the whole theory of plane mirrors. For example, let D represent any other source of waves, its image will be D', as much nearer the mirror than C' as D is than C, and on the same side and at the same distance from the line CC' as is D, and so on for any number of points. It will be observed that, seen from the mirror, D is on the right of C, but the image D' is on the left of C'; hence the image of a system of points by a plane mirror is a repetition of the system, exchanging only right for left. Such a change as this is called a perversion, and the image is said to be a perverted image.

Suppose, now, that the points CD, etc., are either centres of light waves because self-luminous, or centres of disturbance because light waves from other sources fall upon them,

and that AB is an unbounded polished plane like a looking-glass. It is evident that, to the perception of an eye in front of the mirror, the half of the universe behind the mirror is annihilated and replaced by a perverted image of the half in front of it. If the mirror is not unbounded, the problem presented is hardly less simple: as before, we must regard the space behind the plane of the mirror as occupied by the perverted copy of what is in front; but, in addition, we must regard the mirror as a window by which alone we can see into this space. From these elementary considerations it is obvious that we are familiar only with left-handed images of ourselves in a mirror, and that the smallest mirror by which one can see the whole figure has one-half the width and length of the observer, wholly independent of its distance.

A simple and instructive experiment is to look into the angle formed by two vertical strips of looking-glass when held at right angles to each other. In this case, as shown in Figure 3, one mirror forms a perverted image of the region between the two, and the other a perverted image of the first image, so that in the region between the two dotted lines in the figure we have a non-perverted image of the region between the two mirrors. Thus an observer looking into the angle of the mirrors sees a right-handed image of his face, in which any lack of symmetry in the features — in the arrangement of the teeth, for example — becomes very striking to one familiar only with the image as seen in a single mirror, although it is obviously the appearance always presented to his friends. So, too, a printed page held in front of the angle of the two mirrors can be read from left to right in the reflection, instead of from right to left as in the reflection of a single mirror.

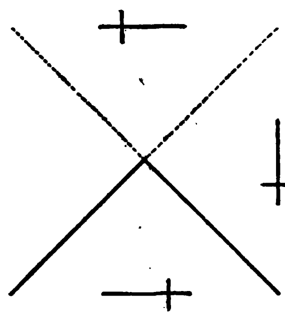

FIGURE 3.

If the reflector is curved, the conditions are far more complicated, and, in general, the reflected waves are no longer spherical, or, in other words, no image is formed. If, however, all portions of the wave fall nearly perpendicularly upon the reflector, the sphericity of the wave is nearly preserved, and an image is formed. If, as in Figure 4, *A*, the reflector *ab* is curved more than the dotted line, which is half-way in respect to curvature between the wave and a flat mirror, a real image of *C* will be formed. On the other hand, if the curvature of the reflector is less than that of the dotted

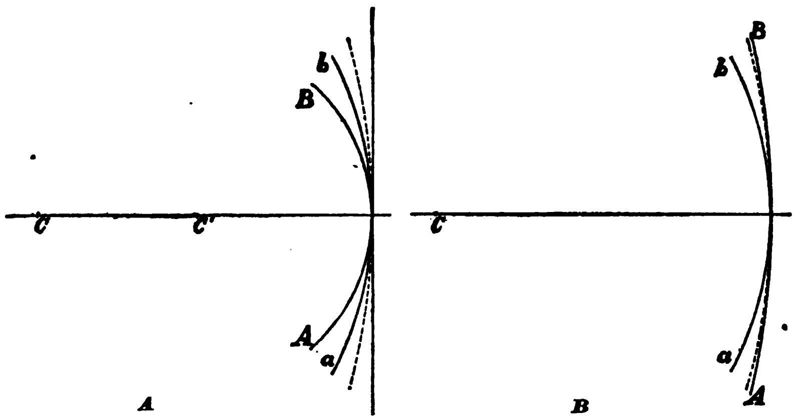

FIGURE 4.

line, as is the case in Figure 4, *B* (which includes that of a reflector curved in the opposite direction or convex toward *C*), we shall have a virtual image of *C* on the opposite side of the mirror. As no use of the properties of curved reflectors other than those here described will be made in subsequent pages, we may content ourselves with the observation (which may be easily proved, however) that, if the mirrors are spherical, and the conditions are such that the image formed is real, this image will not be perverted, but inverted; again, if the image is virtual, it will always be erect and perverted, and also smaller than the object.

Huyghens's explanation of refraction, that is, the change in direction of wave propagation when the waves pass from one medium to another, is as simple as that of reflection, its only assumption being that there is a definite and different velocity of light for each medium. To illustrate his explanation, let AB, in Figure 5, represent the boundary between two media, and C the centre of wave motion in the first medium. We will suppose that the velocity in the second medium is only two-thirds as great as in the first. Then if, as in the case of reflection, we consider the condition of things when a point of the wave o reaches the boundary, we must assume that this point becomes the centre of a wave which is propagated in the second medium, and which at a certain instant, say after the time which is required for the point b of the wave to reach the boundary at p_1, will be found in the circle o' such that oo' equals $\frac{2}{3}$ of bp_1. If this construction is extended to all

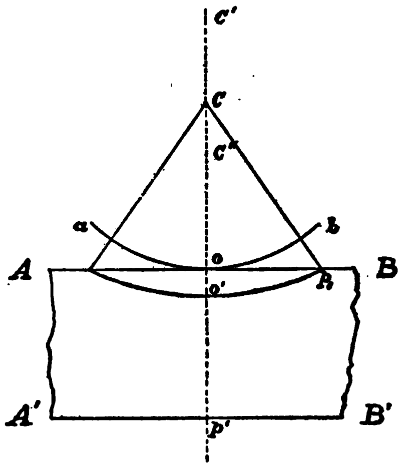

FIGURE 5.

intermediate points of the refracting surface, we shall find that the envelop of all the little waves will be the curve $o'p_1$, the disturbances inside this line being mutually destructive, as proved by the principle of Huyghens. This new wave is not, in general, a spherical wave; hence we do not have an image of C formed; but if we impose the condition that the extent of the wave aob shall be so restricted that it is not *far* removed from a straight line, the refracted wave may be regarded as circular, and its centre C' as the image of C. Since the refracted waves

do not really come from C', but only appear to do so, C' is a *virtual* image of C. The position of the image is fixed by the considerations, first, that it must be on the perpendicular Co because of the obvious symmetry of the construction; and, second, that the distance Co must be three-halves as great as $C'o$ because the curvature of the wave p_1o' is just two-thirds what it would have been without the change in velocity.

If the second medium has another boundary $A'B'$, parallel to the first and at a distance D from the first, the waves in passing through it will have their curvature increased three-halves times, and an image of C' will be formed at C'', which is two-thirds of the distance $C'p'$ from p'; but two-thirds of $C'p'$ are equal to $\frac{2}{3}$ $(C'o + D)$ or to $Co + \frac{2}{3}$ D. This last distance is $\frac{1}{3}$ D less than the distance of the primary centre C from p'. It will be observed that the diminution in distance of the apparent origin of the waves depends only on D, not on the distance of C from the boundary; hence, if C were in contact with AB, its virtual image would still be $\frac{1}{3}$ D nearer $A'B'$, as in the case considered.

If we apply the reasoning to light waves, we have the interesting deductions that, since these do in fact move two-thirds as fast in common glass as in air, any point seen through a plate of glass appears at a distance one-third the thickness of the glass less than its real distance; also that a plate of glass appears only two-thirds its real thickness when one looks through it. As light waves move three-quarters as fast in water as in air, it follows, by exactly the same reasoning, that any vessel of water appears only three-fourths its real depth when we look vertically downward through its surface.

A specially important case in optics is that where the plane surface whence the light emerges is not parallel to the first, or incident surface. This is illustrated in Figure 6. In this case the transmission through AB produces a virtual image of C at C' on the perpendicular Cp, and the transmission through $A'B'$ a virtual image C'' of C' on the perpendicular

line $C'p''$, which is one-third the distance from p' to p'' nearer the latter point than the original centre of waves is to p. Thus it appears that the effect of such a system is not only to bring the image nearer, but also to displace it toward the thin edge of the wedge, if, as in our assumption, the waves travel more slowly in the second medium than in the first. In optics a body which acts upon waves of light in this manner is called a prism.

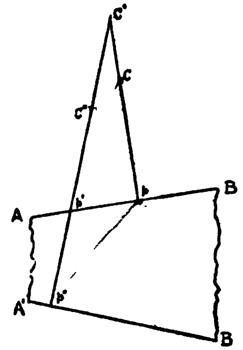

FIGURE 6.

Familiar experience teaches that whenever light passes from one medium to another the transmission is not complete; some of the light is always reflected. Thus the points in the boundary between two media become the centres of two sets of waves, one of which, the reflected light, remains in the original medium, while the other is the refracted wave which we have just been considering. The relative intensities of these two sets are very variable, but the reflected waves are always stronger with increased obliquity, so that for very large angles of incidence the reflected portion which remains in the first medium may be even stronger than the transmitted. There is a singular case of complete reflection when the first medium is that in which the velocity of propagation is the slower of the two. To make this clear, let oo_1, in Figure 7, represent a wave which is moving inside of a block of glass in the direction of the arrow; then, according to the

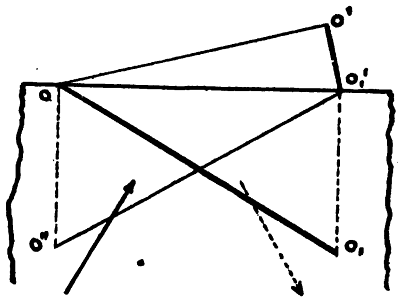

FIGURE 7.

construction which has been used heretofore, o'_1o'' will be the reflected wave moving in the direction of the dotted arrow, and oo' will be the refracted wave in the outer medium, which we may suppose to be air. Now in this figure, if the angle of incidence is increased by a quite small amount, the distance oo' may become just three-halves of oo'_1, in which case the refracted wave would reduce to a length zero. For all larger angles the construction becomes impossible. The angle defined above is called the critical angle, and for it and larger angles there is no refracted light, all being reflected. This phenomenon is called total reflection, and it is readily observed in a smooth glass containing water into which some object — a spoon, for example — is thrust. If one observes from a level considerably below the surface of the water, a feeble image of the object by reflection may be seen, and beyond it the part above the water displaced by refraction; but if the eye is raised so as to approach the level of the free surface of the water, the reflected image becomes as brilliant as the object itself, and nothing can be seen beyond the surface; in short, the surface appears as brilliant and as opaque as a mirror of polished silver.

An experiment to demonstrate the refraction of water waves is very easily devised, and it affords an interesting illustration of the principle involved. It depends upon the fact that the speed of a water wave in shallow water increases with the depth. If a piece of flat board, say half an inch thick, is fastened to the bottom of a large pan, such as is used in baking meats, and then water is poured in until the board is covered to the depth of an eighth of an inch, it will be found that waves produced by immersing and removing a rod at one end of the pan will move with a diminished velocity over the board. If the edge of the board is oblique to the direction of the motion of the waves, the wave follows a new course over the board; also, if the board is wedge-shaped, it will be found that the direction of motion will be permanently altered in passing the shallow portion.

It only remains to consider what would be the effect of a

curved boundary between the two media. In general, the effect would be complex, and no image of the source would be formed; but with the same limitations as those imposed upon the reflection of light waves from a curved surface, namely, that the wave should nowhere make a large angle with this surface, the refracted waves will remain nearly circular, and their centre is then an image of the source. In this case, also, we will suppose the wave velocity one and a half times as great in the first medium as in the second. Then, if C, in Figure 8, represents the source and ApB the curved boundary between the two media, after the disturbance from the source has reached the point p, it will

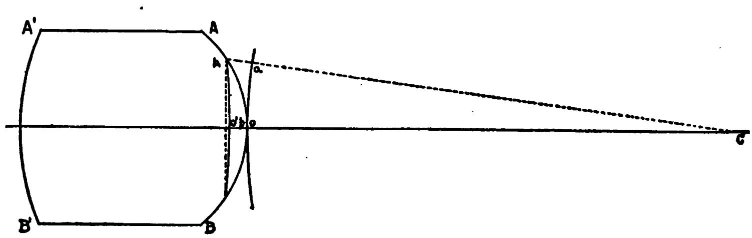

FIGURE 8.

move in the second medium at two-thirds its previous velocity, so that when o_1 reaches p_1 it will have reached o', if po' is made equal to two-thirds of o_1p_1. It is obvious that if the distance po' is less than the distance of p from the dotted straight line in the figure, the wave will be concave in the direction of its motion, and it will move forward to a point which will be a real image of the source C. This is the case in the figure. On the other hand, if the first distance is greater than the second, the wave will be convex in the direction of its progress, and will appear to come from a centre in the first medium, which therefore forms the virtual image of the source. Whether the refracted wave shall be convex or concave is determined solely by the relation of the curvature of the boundary to that of the wave falling upon it, and

to the relative velocities in the two media. We may note the limiting case, namely, if the curvature of the boundary corresponds to that of the wave which falls upon it, the shape of the wave is not altered by refraction, and the virtual image corresponds with the source itself. If the curvature of the boundary is greater than that of the incident wave, the latter is increased, and the virtual image lies between C and p; finally, in all other cases the curvature of the wave is decreased, or even possibly reversed, if the curvature of the boundary is turned the other way. If the effect is the former of these two, the image formed is virtual, but more remote from the boundary than is the source; in the second case the image is real, and on the opposite side of the boundary.

If the second medium is bounded by another curve, so that the waves again enter a medium like the first with accompanying greater velocity of propagation, the change produced in the curvature of the waves will obviously be in the opposite direction from that shown in Figure 8. But if the second boundary is curved in a reversed direction, the effect will be additive, and the total change will be the same in kind but greater in amount.

Applying these results from the theory of wave motion to light, we require no modification in the reasoning, but must bear in mind that light waves spread out in all directions from the centre of disturbance, not merely upon a surface, as in the case of water waves; the boundary between the two media must then be a surface instead of a line. Pieces of transparent substances, such as glass, quartz crystal, rock salt, etc., bounded by polished spherical surfaces, are called *lenses;* they may be regarded as the fundamental elements of all optical instruments. Lenses are divided into two classes: those which are thicker at the centre than at the edge, called positive lenses, and those which have edges thicker than their centres, called negative lenses. The terms "positive" and "negative" are given because the first kind may produce a real image of an object, while the second type never can do so. The shapes of the cross-sections of various lenses of each

kind are shown in Figure 9, the first group being those of positive lenses.

In order to determine the effect upon a wave when passing into a new medium, it is necessary to know the ratio of the wave velocity in the first medium to that in the second. This being given, we can readily determine all the phenomena of refraction, however complicated, by following out the simple construction given in Figure 5. In optics the velocity of light waves in air divided by their velocity in any other substance is called the index of refraction of the substance, a quantity nearly equal to four-thirds for water and to three-halves for common glass. It is an interesting

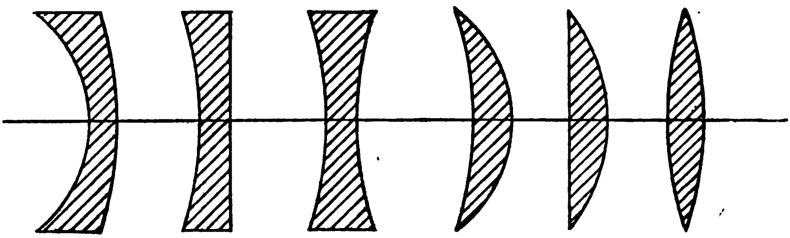

FIGURE 9.

fact that this ratio played as important a rôle in Newton's theory of light as in the wave theory; but it was there inverted, that is, it was absolutely necessary to assume that light travels faster in water than in air. Thus there has been a theoretically simple and conclusive method of establishing one or the other rival theory known to philosophers ever since the second one was proposed. In view of the fact, however, that light moves at the rate of 186,000 miles in a second, the difficulties of an experimental decision between the two theories seemed insurmountable; but, finally, in 1850, Foucault, by means of a rapidly rotating mirror, demonstrated that the deduction from the wave theory of light is alone in accordance with experience. This observation came too late to be of any real philosophical importance, for

the advocates of the wave theory had already fought and won their battle; but the investigation stands high among those which will always challenge the admiration of physicists on account of the skill exhibited by the experimenter, involving, as it did, the measurement of an interval of time less than one one-hundred millionth of a second.

CHAPTER II

OPTICAL INSTRUMENTS

In the last chapter the properties of a lens have been defined in such a manner that its power of forming images, real or imaginary, is implied in the definition. It is essential to establish somewhat more definite notions in this respect.

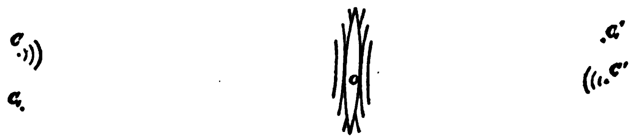

FIGURE 10.

Let o, Figure 10, be a thin double convex lens. Its effect upon a wave from C is to reduce its curvature, or, if sufficiently powerful, to change the direction of its curvature and form a real image of C at C'. The amount by which the lens changes the curvature of the incident wave is called the *power* of the lens. The unit power is that which changes a flat wave to a concave wave of ten inches radius, or, what amounts to exactly the same thing, which can change a convex wave of ten inches radius to a flat one. This is of course a purely arbitrary definition, but it has long been in use, and we shall find it convenient to retain it.

Taking into account the symmetry of both the wave and the bounding surfaces of the lens with respect to the line Co, inspection of the figure will show that the image, wherever formed, must lie somewhere on that line extended. Consider, now, another system of waves from the neighboring

centre C_1. Since it meets the first surface of the lens at o somewhat obliquely, it will have its direction changed; but the refracted wave will meet the second surface, if the lens is thin, at the same angle with which it left the first surface; hence, by a reversal of the refraction, its original direction will be restored, and the only remaining change will be that of curvature. Hence an image of C_1 will be formed at C'_1, on the extended line C_1o. This consideration will hold for any number of points near C, and therefore an image of the region near C is formed near C'. The figure also renders evident the fact that the image is inverted if real, but if virtual, that is, if the centres of the transmitted waves are on the same side of the lens as the original sources, the image is erect. Moreover, it is evident that the ratio of the size of the image to that of the object is the same as the ratio of the distance from the lens to the image to the distance of the lens from the object. In view of the above definition of the power of a lens, we see that a powerful lens produces a small real image of a remote object, but it may produce a large virtual image of a nearer object.

This reasoning concerning the images formed by lenses would require modifications for thick lenses, and for those which have other than symmetrical forms; but the resulting differences may be disregarded in this discussion without affecting the validity of our theoretical conclusions. It is also true that such differences are generally small in practice.

If a piece of white paper is placed at C' perpendicular to oC', and then surrounded by the blackened walls of a box, so that no light except that from the object at C can fall upon it, each point of the paper will appear to the eye bright or dark, according to the brightness of the corresponding point in the object; in short, the surface of the paper will appear as a faithful, though inverted picture of the object. An instrument so constructed is called a camera obscura. Used formerly only as a scientific toy, or occasionally as an aid in drawing the outlines of an object, it has become, since the invention of the various methods of

fixing an image on a sensitive plate, one of the most important of optical instruments. Since it is necessary to have the surface upon which the picture is formed at a distance from the lens depending upon the distance of the object to be pictured, the sides of the photographer's camera box are made extensible, like the flexible portion of a bellows, and the adjustment is secured by aid of the eye before the sensitive plate is introduced.

Optically the eye itself is simply a camera obscura. It is figured in section on page 156, Figure 40, and in the eighth chapter a somewhat extended description of its structure and action is given. Here it is only necessary to point out the one striking difference, from an optical standpoint, between it and the photographer's camera, namely, the means for adapting the apparatus to varying distances of the object observed. Instead of modifying the distance of the screen from the lens, the power of the lens itself is changed by changing its thickness in the middle. But as very many eyes have either lost this power in part, or have never possessed it in perfection, we shall assume hereafter, when an experiment is described which is to be performed with the eye, either that the eye is a normal one or that it is rendered equivalent to a normal eye by the employment of the proper spectacle lens.

This brief discussion of the optical properties of the eye serves to define the conditions of distinct vision of an object or of an image of an object. If object or image is within the range of distinct vision, that is, in front of the eye and at any distance greater than five inches, it can be distinctly seen if sufficiently large and bright. In those cases where a definite value for the distance of distinct vision is necessary for comparisons, ten inches is arbitrarily chosen. For example, if we say that the magnifying power of a microscope is one hundred times, we mean that the object seen through the instrument appears a hundred times larger in every dimension than it would if held at a distance of ten inches in front of the unassisted eye.

If an object is too small to be seen with sufficient distinct-

ness, it may be made to appear larger by bringing it nearer to the eye, so that at half the distance it appears twice as large, and so on for other approximations. But if this means of increasing the apparent size of an object is carried too far, the power of the eye becomes insufficient to change convex wave-surfaces originating at the object to concave ones having their centres on the retina. Under such circumstances we must add to the power of the eye lens by employing a positive lens close to the eye. A lens so used is called a simple *microscope*. If the power of this lens is considerable, that is, sufficient to convert waves from a point one inch or less in front of it into flat waves, it is easy to show that increased size of the image formed on the retina is proportional to the power of the lens. It is this relation which gives rise to the term "power" to designate that particular constant of a lens.

Since the eye is optically a camera obscura, it follows that the image depicted upon the retina is inverted with respect to the object. Why this inversion does not appear in the sensation is a question of psychology rather than of optics. It is sufficient for our ends to recognize that an inverted image on the retina corresponds to the sensation of an erect object, and *vice versa*.

The preceding discussion of the optics of the camera and of the eye is important on account of the fact that it practically embraces the purely geometrical theory of all optical instruments. For example, the magic lantern and the solar microscope are cameras in which the object is quite close to the lens and the image remote. The only feature besides this, and one which is common to both, is an arrangement by which the object can be very strongly illuminated. Then, also, the telescope and the compound microscope are each a combination of the camera and the eye, the power of the latter being usually increased by a lens which receives the name of eyepiece, or ocular. These we shall consider somewhat more at length after a description of the optical principles of the opera glass.

The Galilean telescope, whether considered from its his-

tory, its achievements, or its theory, is one of the most interesting of all optical instruments. Galileo learned in 1609, while visiting Venice, that a wonderful instrument had been invented the preceding year in Holland, which would enable an observer to see a distant object with the same distinctness as if it were only at a small fraction of its real distance. It required but little time for the greatest physicist of his age to master the problem thus suggested to his mind, and after his return to Padua, where he held the position of professor of mathematics in the famous university of that city, he set himself earnestly at work making telescopes. Such was his success that in August of the same year he sent to the Venetian Senate a more perfect instrument than they had been able to procure from Holland; and in January of the next year, by means of a telescope magnifying thirty times, he discovered the four large satellites of Jupiter. This brilliant discovery was followed by that of the mountains in the moon; of the variable phases of Venus, which established the Copernican theory of the solar system as incontestible; of the true nature of the Milky Way, and by many others of less philosophical importance. Though Galileo did not change the character of the telescope as it was known to its discoverer in Holland, he made it much more perfect, and, above all, made the first and most fertile application of the instrument to increase the bounds of human knowledge, so that it is inevitable that his name should be indissolubly connected with the instrument.

Considering the enormous interest excited throughout intellectual Europe by the invention of the telescope, it seems surprising that its early history is so confused. Less than two years after it was first heard of, a discovery — perhaps the greatest in the domain of natural philosophy for a thousand years — had been made by its means. Notwithstanding these facts, the three contemporary or nearly contemporary investigators assign the honor of its invention to three different persons, and if we should write out the names of all those to whom more modern writers have attributed the

invention the list would be a long one. The surprise will be somewhat lessened, however, if we consider the task before a historian in the next century who undertakes to justly apportion the honor of the invention of the telephone among its numerous claimants. The analogy, though suggested by the evident fact that the telephone is to hearing just what the telescope is to sight, could only be close if the future historian were deprived of all but verbal descriptions, and if contemporary models and diagrams were wholly wanting. Under such conditions it is difficult to believe that the historian would easily escape antedating the discovery of the telephone proper on account of descriptions, generally imperfect, of the acoustic telephone. But this would fairly represent the condition of the material at the command of an investigator of the present day in a question of science of the early part of the seventeenth century. No wonder, then, that the invention has been attributed to Archimedes, to Roger Bacon, to Porta, and to many others who have written on optics; but to find the name of Satan in the list is certainly surprising. Still we read that a very learned man of the seventeenth century, named Arias Montanus, finds in the fourth chapter of Matthew, eighth verse, evidence that Satan possessed, and probably invented, a telescope; otherwise how could he have "shown Him all the kingdoms of the world, and the glory of them"? [1]

It seems to be well established now, however, that Franz Lippershey, or Lippersheim, a spectacle-maker at Middleburg, was the real inventor of the telescope, and that Galileo's first telescope, avowedly suggested by news of the Hollander's achievement, was an independent invention.

The Galilean telescope consists of a negative lens *a*, Figure 11, close to the eye, and a positive lens *b* at a certain distance from the first, depending upon its power. For the sake of simplicity we shall suppose that the lens *a* just neutralizes the lens of the eye. This supposition entails no

[1] The early history of the telescope is admirably treated in Poggendorf's 'Geschichte der Physik,' from which some of the foregoing statements are taken.

loss of generality in our conclusions, and it is very nearly in accordance with practice in the instrument as it now survives. In this case the waves suffer no change of curvature in passing into the eye, and if, after passing it, the lens *b* has

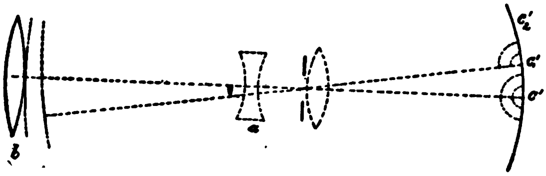

FIGURE 11.

such a position that waves from a distant source have their centres of curvature on the retina, the conditions of distinct vision are met. Thus the effect of the instrument is merely to increase the virtual size of the eye, and with it, as appears from the theory of the camera, the size of the image on the retina in the same ratio. In one particular only does this apparatus — Galilean telescope and eye — differ from an enlarged eye, and in that one difference lies the limitation of this form of telescope. In the eye the iris, which limits the portion of the lens upon which light waves fall, is close to the lens itself and remote from the retina; in this magnified eye, however, the iris is relatively near the retina and remote from the lens. From this it follows that the light coming to different points of the image passes through different portions of the lens *b*. This is evident from the diagram, for the waves which pass through the pupil of the eye and converge to c' come from that portion of the objective *b* indicated by the unbroken part of the wave, while the light which goes to form the point c'_1 comes from another portion similarly indicated. A point in the object which gives rise to the waves that, after passing through the objective, have their centre at c'_2, cannot be seen at all, since these waves are stopped by the iris. Thus the *extent* of that portion of an object which can be seen with such a telescope depends upon the diameter of the objective and of the pupil of the

eye. If the magnifying power is considerable, this extent is very small with moderate-sized lenses, and large lenses are not only expensive but they are impracticable for short telescopes, because the wave-surfaces transmitted by them are no longer spherical. An instrument having this defect is said to have a small field. But even this field is not illuminated quite to the edge, for it is evident that points on the retina between c'_1 and c'_2 receive light from an area of the objective less than that which yields light to points nearer the axis of the eye; and the more remote from c'_1, where a portion of the light comes from the extreme margin of the objective, the greater this disparity. Hence the field presents to the eye an outer marginal circle of indistinct vision. For these reasons this type of telescope is adapted only for use where very moderate magnification, say from two to five or six times, is required, and where extent of field may be sacrificed to compactness and lightness. In opera glasses, with a magnification ordinarily of three or three and a half times, it serves well, and finds a rival only in the ingenious prism telescope which will be described later. It is also used with the sextant, where lightness and shortness are both very desirable; but in almost all other cases it is replaced by a type proposed by the famous astronomer Keppler, in 1611.

Keppler's telescope, ordinarily called the astronomical telescope, may be readily understood from a consideration of the camera obscura, as illustrated in Figure 10. Here the objective forms an image of distant objects on the screen. Now imagine the eye just at the centre of the lens; if turned toward the object, this would appear of a certain magnitude; if toward the screen, the inverted image there depicted would appear of the same magnitude as the object itself, as follows from the principle explained on page 20. If now the eye is brought nearer to the screen, the image will appear larger than the object in the inverse ratio of the distance of the eye and of the lens from the image. This illustrates the general principle involved, although in practice the screen is omitted and the eye is placed in the prolonga-

tion of the axis of the lens, because this permits a larger portion of the total light transmitted by the objective to reach the retina.

It thus appears that in the astronomical telescope the object is seen inverted, a matter of no importance in astronomical observations, and that the magnifying power is equal to the distance from the objective to the image divided by the distance from the image to the eye. If, in order to secure greater magnification, the eye is brought very near the image, its power must be increased by a lens used as a simple microscope, just as in the case of a real object which is too near for distinct seeing. Such a lens, or, preferably, system of lenses, is called the eyepiece, or ocular. The magnification then becomes the power of the ocular divided by the power of the objective.

The figure makes evident that in this variety of telescope, unlike the one previously considered, light from every portion of the objective goes to every point of the image on the retina; hence the size of the field does not depend at all upon the size of the objective, nor is there any variation of brightness toward its margin. This is an advantage of great moment in powerful instruments.

The compound microscope may be conveniently regarded as an inverting telescope adjusted for an object very near the objective. The geometrical principles involved in the

FIGURE 12.

construction are illustrated in Figure 12, where *b* is the objective which forms an image of the object *c* at *c′*, which, in turn, is observed by means of an ocular *a*. The size of the

image is to that of the object as the distance bc' is to bc, or very nearly the power of the objective multiplied by the length of the tube. This image, again, is magnified by the ocular in the ratio of the power of the ocular; hence the whole magnification is equal to the product of the power of the ocular by the length of the tube and by the power of the objective. It is evident from the figure that such an instrument shows an inverted image of the object. The compound microscope is in no respect theoretically superior to the simple microscope, but it is impracticable to make simple microscopes of very great power, say with a magnification of much more than 250, because of the extreme minuteness of the requisite lenses.

If the ocular of the astronomical telescope is replaced by a compound microscope, we shall see an inverted image of the image formed by the objective, that is, an erect image of the distant object. This constitutes the terrestrial telescope, or spyglass. It is necessarily longer than the astronomical telescope and less perfect, because it is subject to the unavoidable defects of a larger number of lenses with their accompanying loss of light. It is employed with good reason for the ordinary purposes of a spyglass, but its customary use in surveyors' instruments is to be deprecated.

The prism telescope, which was invented nearly a half century ago by Porro, an Italian engineer, has recently been perfected mechanically, and is rapidly replacing the simpler Galilean construction for powers from four to twelve. Its principle and construction may be understood from a study of Figure 13, in which b represents an objective, and i_1 an inverted image of the distant object to be observed. If a rectangular prism p is placed so that one-half of it intercepts the light from b, an image i'_1 would be formed by total reflection from the shorter face of the prism at the place indicated in the figure were the light not again reflected from the second surface so that a real image is formed at i_2. This last image is of course an unperverted repetition of i_1, because formed by a rectangular mirror, but if examined from the

side *a*, as might be accomplished either by receiving the light upon a white screen or by reversing the direction of the light by a mirror, it would appear to be perverted in the

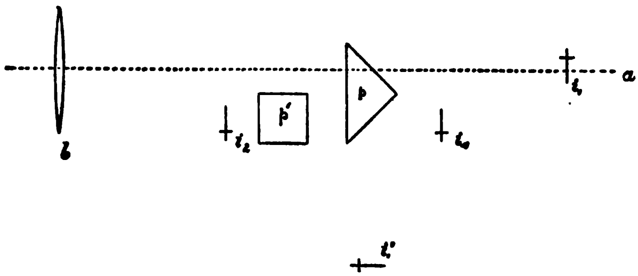

FIGURE 13.

plane of the diagram. If now a second prism p', quite like the first, but with its base at right angles to the plane of the paper instead of parallel with it, receives the light from p on the lower part of its largest surface, it will form, after two total reflections, an image i_4 above the plane of the figure. Thus by the four successive reflections the final image is a complete inversion of i_1, and the light is also restored to its original direction, whence the image can be observed by an ocular with the eye turned toward the object. The advantages of this construction are sufficiently obvious. First, the objective and ocular are essentially those of the astronomical telescope, and it retains, therefore, the moderate ratio of aperture to power and the relatively large field. Second, the instrument is very much shortened, so that it is rendered conveniently portable. On the other hand, it has defects which are far from immaterial. The loss of light from the reflections from the larger faces of the two prisms and by absorption in their material is by no means inconsiderable, and, as might be supposed, objects appear notably less bright than through a Galilean telescope. Then the awkward form necessitated by the offset in the course of the light introduces structural difficulties which doubtless explain

the very long time during which the invention remained undeveloped.

This completes our review of the purely geometrical principles involved in all the more important optical instruments. It will be observed that the magnification in each case depends very simply upon the powers of the lenses which enter into the several constructions; but when we come to discuss the efficiency of such instruments, we shall find that these powers play a wholly insignificant rôle; in short, that the effective diameters alone are the measures of optical efficiency.

CHAPTER III

PHENOMENA OF LIMITED WAVE-SURFACES—INTERFERENCE—WAVELENGTHS OF LIGHT

HERETOFORE we have considered only those phenomena which necessarily follow from the rectilinear propagation of waves, itself a consequence of Huyghens's principle, and those dependent on the varying velocities in different media. We now enter upon a consideration of the consequences entailed by limiting the extent of the wave-front; and we shall see that these consequences, very interesting in themselves, as well as of extensive application to the explanation of phenomena of every-day observation, are connected by simple laws with the absolute value of the lengths of the waves.

Suppose Figure 14 to represent a wave-front limited by the screen *ab*, moving forward toward the centre *p*. The point *p* is thus, according to definition, the image of the point from which the wave took its origin. According to the principle of Huyghens, each point of the wave-front *ab* must be regarded as a centre of disturbances which are propagated in all directions through the medium. Consider the condition at p_1, so chosen that its distance from *a* is one-half a wavelength greater than its distance from *b*, that is, so that *aa′* equals a half wavelength. Here a disturbance setting out from *a* will reach p_1 at the same instant as one starting from *b* one-half a period later, in short, the motion derived from *a* will be exactly equal and opposite to that from *b*. Thus the effects of these two pairs of points in the wave *ab* will perfectly neutralize each other. But these two

points are the only ones in the limited wave-front *ab* so related to p_1; hence the effect of other pairs of points, which we might imagine chosen in the wave-front symmetrically placed with respect to its middle point, will only partially

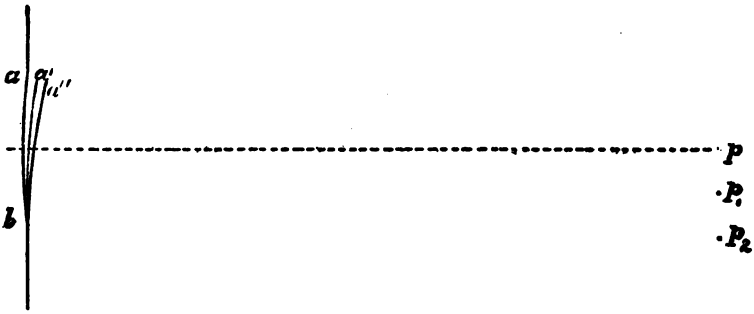

FIGURE 14.

counteract each other. For points between p and p_1 the mutual destruction of the elementary waves is even less complete, as there is no pair of points which wholly destroys each other's effect. Take now the point p_2, such that the distance from a is a whole wavelength greater than its distance from b. In this case we see that the disturbance from the middle point of the wave-front will be a half wavelength behind that from b, and will thus destroy its effect; but for every point between b and the middle of the wave-front a corresponding point between the middle and a can be found which is a half wavelength further from p_2; hence the effects of all the elementary waves at p_2 would be nil, and there the medium would be perfectly quiescent. If we consider points more remote from p, we shall find, by the extension of the same reasoning, that in general there will be motion due to the effect of the bounded wave-front, except where the difference of the distance from a and b is a whole number of wavelengths. When this difference is an odd number of half wavelengths, the disturbance is greater than at any closely neighboring point, because the conditions of mutual destruction are most widely departed from; on the other hand, it is obvious that the absolute

value of the disturbance decreases very rapidly in leaving the position of the geometrical image p, because a larger and larger number of pairs of mutually destructive centres can be found. The light which is found outside the geometrical image is called diffracted light, and a large class of analogous phenomena are embraced under the general term of diffraction.

The conclusion from this study is, that a limited concave wave-front forms, not a simple image at its geometrical centre, as we have assumed heretofore, but a series of images of which the middle one corresponds in place with the geometrical image, and is by far the strongest, while it is symmetrically flanked on either side by secondary images rapidly diminishing in intensity. But we are able to infer far more than this. It is evident from the figure that the distance pp_2, which is half the diameter of the central image, bears the same ratio to the distance ap as aa'', or one wavelength, does to the distance ab.

If the waves are those of light and the wave-front is a spherical surface bounded by a circular aperture of which ab may be taken to represent a diameter, we have the most common and most interesting case of optics, for it is practically that of all optical instruments. In this case, although the phenomenon is more complex in its numerical relations, it does not differ in kind. The image proper becomes a circular area instead of a point, and the secondary images are concentric circles, as we should expect; but the distances from the centre to the various rings are slightly greater than the values derived from the simple treatment of the last paragraph. For example, the distance from the centre of the primary image to the first ring is 1.2 times as great as from p to p_2 in the figure. It is this relation that we shall find of importance when we come to the consideration of the limitations of optical instruments.

Since the foregoing considerations are perfectly general and applicable to all kinds of waves, the question naturally arises, Why do we so rarely notice such extraordinary

phenomena in the numerous examples of wave motions about us? The answer to this question will be discovered in analyzing the somewhat rigid conditions implied by the reasoning connected with Figure 14. First of all, in order to secure regular phenomena about the region p, we must have a series of waves of uniform length, since the distance from the primary to the secondary image depends upon this length. But upon the surface of water the waves are generally of the utmost irregularity, nor do they ordinarily converge toward a centre, so it is clear that highly artificial and unusual conditions must exist in order to make these particular effects strikingly obvious; if these exceptional conditions are observed, however, there is no difficulty in verifying the results of the theory.

Again, if the opening through which the waves come is less than a wavelength, there is no point, such as p_2, whose distance from one edge of the aperture is a whole wavelength greater than that from the other; consequently the most striking peculiarity of the effect, namely, the regions of quiescence, is wholly wanting. In the case of sound the average lengths of the waves corresponding to the voice of a man is not far from eight feet and of a woman's voice about half as great; hence we ought to have apertures a number of feet across in order to produce the required effect. But as soon as our apertures become as great as this, they are of the same order of magnitude as are the enclosures in which our observations are made, so that waves reflected from the surrounding walls entirely mask the effects sought, at least to one not guided by theory in his search. Thus it is not surprising that this particular class of phenomena in sound waves failed of recognition until after their discovery in the case of light.

Finally, if the aperture is many times larger than the length of the wave, the nearer and stronger images will lie so close to the primary as to escape, perhaps, our powers of perception. This is generally the case with light where the waves are very short.

Since to our eyes a star appears as a point of light, not as a spot of light surrounded by a series of concentric circles, notwithstanding that the waves from the source are in fact bounded by the circular inner edge of the iris, we must conclude that light waves are very many times shorter than the diameter of the pupil of the eye, or, in other words, that very many wavelengths would be comprised in the length of about one-eighth of an inch. According to the theory, the separation of the images increases directly as the aperture decreases; consequently, if we restrict the aperture of the eye, we ought finally to secure a perceptible separation of the secondary images. Thus, if we look through a needle hole in a card at a very bright point, such as a distant electric light, or, more conveniently, a bright bead or thermometer bulb in sunshine, the central disk and luminous surrounding rings become obvious at once. Moreover, in accordance with theory, the smaller the hole in the card the larger the disk and its concentric rings, though of course fainter as less light is allowed to enter the eye. If the luminous point, which we may conveniently call an artificial star, is very brilliant, two or three or even more rings may be seen; but if a less brilliant source is employed, the outer rings, which very rapidly decrease in brightness, will be too faint to be seen.

The modifications in the image produced by a concave wave-surface passing through two small holes in an opaque screen are particularly interesting, because we shall find that they yield a ready means of measuring the value of the length of the light waves.

Suppose *ab* and *a′b′*, Figure 15, to represent two circular holes through which the wave-front whose centre is at *p* passes. Then the light which passes through *ab* forms an image in the region about *p*, consisting of a central disk and concentric rings, as we have seen; so, too, the light which passes through *a′b′* forms a similar image at the same place. If the waves which pass through the former opening had no definite relations to those which pass through the latter, as

would be the case, for example, if the waves through *ab* came from one artificial star and those through *a′b′* from another, the effect would be simply to make the image twice as bright, since twice as much light falls upon that portion of the screen. But if the waves at the two apertures are congruent, which would be the case if they came from the same source and had been subject to the same conditions before reaching the screen, the image will be profoundly changed. In this case we should still find the disk and concentric rings, but they would be crossed by a series of fine, dark, nearly straight lines perpendicular in direction to the line joining the centres of the holes. That this must be

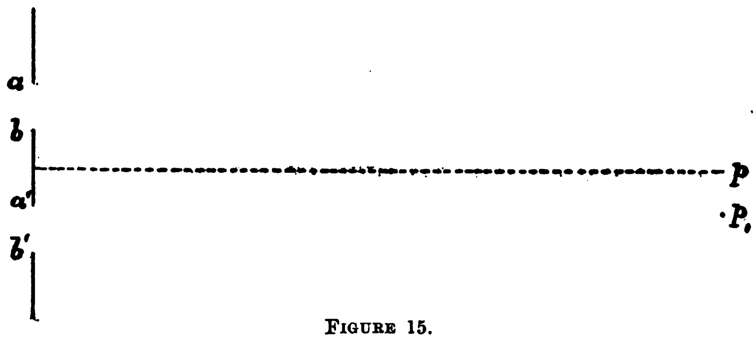

FIGURE 15.

so follows from simple considerations of the principles involved. The distance from the centre to the first dark ring, which we may call the radius of the primary image, is, as has been shown in the discussion connected with Figure 14, nearly 1.2 times the distance of a wavelength multiplied by the ratio of *ap* to *ab*. But starting from *p*, long before we come to the point p_1, we come to a region *p′*, such that the difference of its distance from *a* and *a′* is half a wavelength, and hence the effect produced at *p′* by the portion of the wave *a* is destroyed by that produced by *a′*. If, however, the centre *a* is thus neutralized by the centre *a′*, the effect of every other point within the region *ab* will be neutralized by the corresponding point in *a′b′*; consequently we shall

have no light at p'. But this reasoning is equally applicable to a point p'', which is at three-halves wavelengths greater distance from a' than from a, and also when this difference is any *odd* number of half wavelengths. Hence there will be a series of dark lines such as described. It will be observed that the diameter of the image is inversely as the diameter of the apertures, but that the separation of the dark lines is inversely as the distance between the centres of the holes, the two features of the phenomenon thus depending on entirely different elements. This interesting phenomenon can be easily seen and the theory verified by making two needle holes in a card at a distance considerably less than the diameter of the pupil of the eye, and looking through them at an artificial star. Figure 16 is a photographic reproduction of this appearance in which an ordinary camera replaced the eye and a sensitive plate the retina.

FIGURE 16.

The chief interest of the experiment, as noted above, lies in the fact that it enables us to measure the length of light waves with surprising precision with no other apparatus than a finely divided scale, say to hundredths of an inch, which may be obtained at any hardware store. Even a carpenter's rule, divided to sixteenths of an inch, will enable us to measure waves of light by this means with more ease than we can measure water waves to the same degree of precision. As this value is the fundamental unit of the whole theory of light and of nearly every phenomenon which is studied in the following pages, the reader is strongly recommended to make the measurement for himself.

To do so, make a series of pairs of fine holes in a card with a needle and look through them at an artificial star. If the pair of holes is separated by an interval of a twelfth of an inch or more, no lines across the disk-image will be seen; if the interval is a twentieth of an inch or less, the lines become very distinct. Select a pair of holes at such a dis-

tance that the interference lines can be *just* seen, and measure the distance between the holes with the scale. It will be found easy to measure this interval to much closer than one two-hundredth part of an inch, if the scale is divided to hundredths of an inch, and to a hundred-and-sixtieth part by estimating tenths of the interval of an ordinary carpenter's rule. Call this distance D. We require now to determine the ratio of the distance $p'p''$ (or its equal pp_1) to the distance ap, referring to Figure 15. Neither of these quantities can be determined readily by itself, but the ratio can be easily found as follows: Draw a series of parallel lines on white paper in ink, making the width of the lines approximately equal to the distance between them. Fasten this paper to the wall and find the distance at which they can be just seen as separate lines, the eye being aided by the proper spectacle glass if necessary. The ratio of the distance apart of the lines to this distance is the ratio sought, since by supposition the system of interference lines was also just visible as separate lines. Finally, as appears from Figure 15, the wavelength of light is D times this ratio.

For examples we may cite the following: It was found by an observer that, looking through certain pairs of holes at a thermometer bulb in sunshine, no lines could be seen through the first, very distinct lines through the second, and the finest possible lines through the third. A measurement of the intervals gave 0.08, 0.05, and 0.065 of an inch, respectively. For that observer's eye, then, D equals 0.065 of an inch. Then five parallel lines were drawn on paper which, it was found, could be just seen as separate lines at a distance of twenty-seven feet. The measurement of the group of lines showed their width to be 0.48 of an inch, and consequently 0.12 of an inch as the interval from one line to the next. The required ratio is therefore 0.12 of an inch to 27 feet, which is 2700, and the wavelength of light is 0.067 of an inch divided by 2700, which is almost exactly 1/40000 of an inch. A second observation, in which only a carpenter's rule was employed and all the quantities involved

were redetermined, gave 1.05/16 of an inch for D, and 1/20 of an inch for the distance between the lines on the paper which could be seen as vanishingly fine from a distance of twelve feet. These data by the same process of reduction give 1/2880 of the distance between the two holes for the length of the wave, or 1/45000 of an inch. The estimated value of D means, of course, that the distance from one hole to the other is a little greater than one-sixteenth of an inch, but not so much as a tenth greater.

Of these values, not very discordant in themselves, the second is the better, and is very close indeed to the true value, which for white light may be regarded as equal to 1/45200 of an inch. In this better determination the paper with parallel lines to find the resolving power of the eye was put in full sunshine, thus approximating more closely to the condition of very bright and very dark lines, such as those of the interference phenomenon.

If three holes are made in the card, there must be a set of dark lines for each pair, that is, three sets of lines. Thus the general effect will be that of three groups of lines crossing each other at the same angles as those of the triangle formed by the centres of the three holes. This is not true with absolute exactness, because a region where the effects of two of the apertures destroy each other is still illuminated by light from the third, so that extremely beautiful and complicated patterns may occur, although the general configuration is given by the simple rule of a set of interference lines for each pair of holes. For example, three holes at equal distances from each other yield a series of hexagonal images arranged as are the cells in a honeycomb, although to the eye it would appear simply as three systems of equidistant lines crossing the disk at mutual angles of 60°. More important is the deduction which we can make from the indifference as regards the position of the holes before the eye, provided only that the light from no one of them is cut off from the retina by the iris. That this position is inessential follows from the discussion of Figure 15, where the

iris is quite left out of the question. Hence we might have two systems of holes of exactly the same size and configuration so far apart that light from one system would not visibly modify that from the other; in this case the effect would be only to double the quantity of light which forms the image. But it is not difficult to see that the second system of holes need not be remote from the first if there are no new distances introduced except multiples of the original distances. Hence, if a piece of perforated cardboard, such as is used for worsted embroidery, is held before the eye while looking at the artificial star, the effect, though much brighter and the separate images smaller, is the same in kind as though only four of the holes were transparent. The phenomena presented by small luminous points seen through fine and regularly woven fabrics, such as silk, lawn, bolting-cloths, wire cloth, etc., are of this kind. Many feathers will exhibit beautiful effects in this way. A modification of the phenomenon may be made by holding the perforated cardboard in front of the objective of a telescope directed toward a bright star. In this case we virtually increase the dimensions of the eye and can use correspondingly coarse structures in the perforated screen. Many of these figures are of surprising beauty, and in all of those where the image is bright and sufficiently large, brilliant colors are seen. The explanation of these colors we shall find in the next chapter.

CHAPTER IV

DISPERSION — CHROMATIC EFFECTS OF DIFFERING WAVELENGTHS — COLORS OF THIN PLATES

HERETOFORE we have tacitly assumed that the waves given out by a luminous point are all alike, under which assumption we have explained by means of Huyghens's principle the phenomena of reflection, refraction, and interference. But in general waves of a very great range of length are emitted by such a point. This fact changes nothing in the reasoning concerning reflection from large surfaces; but interference, which depends upon the length of the waves, becomes obviously more complicated in such cases; and, what is not obvious from that which goes before, the phenomena of refraction are also modified. To show this, vary the experiment described on page 13 with a wedge-shaped glass, or prism. If sunlight is transmitted through such a body and received on a distant white screen, it will be found that instead of a white spot marking the region where the deflected light falls, we shall have a brilliantly colored strip, the end nearest the place where the undeviated light would fall being red, and that most remote from this point violet. The intermediate colors, taken in order from the red, are orange, yellow, green, green-blue, blue. The change from one of these hues to the next is absolutely continuous, so that the number of colors is limited only by the number of names at our command for designating them. Since the change in direction of propagation of the waves depends only on their less velocity in glass than in air, we conclude that those waves which produce the sensation of red move less slowly than those which give rise to the sensation of orange,

and so on. This separation according to wave slowness in any refracting substance is called *dispersion*, and the resulting colored strip is called a *spectrum*. Sir Isaac Newton was the first one to investigate this phenomenon in a scientific manner, and to fix its terminology, though he used the color names green, blue, indigo-blue, and violet, instead of green, green-blue, blue, and violet, which modern writers have found more appropriate. Newton's most important discovery was that after the light is thus modified any one color suffers no further change by passing through another prism. His conclusion was that ordinary white light is compound and made up of an indefinite number of hues, of which seven are recognized by familiar names. He confirmed this deduction by showing that if these differently colored lights were united, either by allowing them to fall upon a concave mirror and reflecting them to a common point, or by passing the decomposed light through a similar prism turned in an opposite direction, the result would be white light like that from the original source. Thus both analysis and synthesis were employed at his hands to demonstrate the composite nature of white light.

Newton also found that like prisms of different substances would produce quite dissimilar amounts of dispersion; but, fortunately for the development of practical optics, his conclusion, that the phenomenon of dispersion (which should be looked upon as a secondary phenomenon of refraction on account of its relative minuteness) increases directly as the refractive power, was an error.

From the Newtonian standpoint the explanation of dispersion was found in the assumption of as many different kinds of small corpuscles as there are different hues; in the wave theory the explanation can be found only in attributing the differences of velocity in the material of the prism and the corresponding colors to differing wavelengths. That the latter explanation is adequate and in accordance with the facts, we may prove by means of the perforated card and the artificial star. This may be done in various ways. We

may form a spectrum by a prism held in the sunlight, and observe the artificial star produced when the thermometer bulb is moved from one end to the other of the spectrum. If the star which is of the color of that portion of the spectrum in which it is immersed, is observed through the two needle holes, it will be seen at once that the interference lines are furthest apart in the red and continuously approach as the star changes successively to orange, yellow, green, etc. Or we may place the thermometer bulb in the sunshine and look at it through a prism; the artificial star will then appear as a fine linear spectrum. If, then, the screen with the two needle holes is placed close to the eye, between it and the prism, with the line joining the holes at right angles to the spectrum, the latter will appear broadened and traversed by a series of fine longitudinal lines, which are the interference lines. It will be easy to recognize that these lines are most widely separated at the red end of the spectrum. The second method, though not quite as perspicuous as the first, has the advantage of requiring only a very small prism; the first requires a large prism in order to have the star sufficiently bright.

A third method, which enables us to dispense altogether with a prism, though otherwise inferior, is to employ colored glasses in conjunction with the pierced card. Thus one observer found that looking through the card and red glass he could see the interference lines when the holes were one-twelfth of an inch apart, with yellow glass it required a distance not much greater than one-sixteenth, and with blue glass he found it necessary to have the holes within one-twentieth of an inch of each other.

The conclusions from these experiments are, first, that the velocity of propagation of light waves in glass decreases continuously with decrease of wavelength; second, that waves of different lengths falling on the retina produce different color sensations, the longest waves awakening the sensation of red and the shortest that of violet; and third, as appears from the quantities given in the third experiment, that the

violet waves cannot be very much greater or very much less than half the length of the red waves. As regards this last conclusion, we may add that it is impossible to state the limiting values of the waves which can stimulate the retina, but without extraordinary precautions we may fairly assume 1/27000 of an inch for the extreme red and 1/52000 of an inch for the extreme violet; thus the ratio of the longest to the shortest is about 1.94 to 1.

To this enlarged view of the nature of white light it is easy to adapt the conclusions of the preceding chapters. As regards the phenomena of reflection, nothing is to be added to what appears in the first chapter; but refraction at a plane surface, as in Figure 5, forms, instead of a virtual image of the source, a virtual spectrum of which the violet end is nearer the refracting surface. The action of a plate is to form a virtual spectrum with the violet end nearer the plate, but this effect can be seen only in the case of a very thick plate, and ordinarily when the object is looked at obliquely. Thus a white pebble seen vertically downward through deep water still appears white, but if seen away from the plumb-line, colors of the spectrum become very obvious. So a prism instead of forming an image of the source displaced toward the thin edge of the prism and approached by a fraction of the thickness of the prism, as appears from the consideration of Figure 6, in reality forms a spectrum with the blue end more displaced than the red end, and also brought a very little nearer.

The description of interference phenomena produced by limited wave-surfaces is quite easily extended to include the color effects due to varying wavelengths. It is only necessary to remember that the images thus produced have dimensions in direct ratio to the wavelength of the light which goes to form them. Accordingly, the image of a white artificial star consists of a disk of which the centre is white, since all waves are there represented, but with a red or orange margin; the rings immediately surrounding the disk are blue on the inner side and red externally. Though this

is the true description of the image, it is impossible to recognize it as such because of the limits of our perceptive faculties, for if the aperture is made large so as to render the colors brilliant, the disk and rings will be too small to be perceived in their details; on the other hand, if the opening is very small, so that the various features of the image are large enough to be obvious, the light will be so faint that the colors are unrecognizable, just as we are unable to name colors in even quite bright moonlight. With two or more apertures, however, we may succeed better. In the case of two holes we may describe the resulting image as composed of an indefinite number of superimposed images, all having the same centre, but of sizes increasing regularly with the color from violet to red. Hence the middle band would be white, flanked on either side by bands of which the inner edge is blue and the outer red. As we go outward from the centre, the chromatic separation will become greater and greater, until finally we reach a point beyond which every color, though of course not every wavelength, will be represented at all points. In all such regions the image will appear white, and the immediate effects of interference will vanish. This limit in the case of white light is found to be reached in the region of the eighth or ninth band, so that in white light we can never see more than seventeen or nineteen bands, even under favorable circumstances, though with light of a single color it is sometimes possible to see many thousands.

In the experiment with two apertures the chief difficulty in seeing the colors lies in the fineness of the lines and the faintness of the illumination, as indicated above; but with holes a thirtieth of an inch apart or less, the lines are sufficiently coarse, though if many bands are to be seen the holes must be very small. We have learned, however, that greatly increasing the number of apertures, provided that they are systematically arranged, produces little change except increased brightness; hence by this means the color effects may be made very marked. An artificial star seen through fine,

uniformly woven silk or through the web of a uniform feather will show very beautifully colored figures. As both nature and art present us with numerous examples of such structures, we could hardly fail to be familiar with this class of phenomena were it not that we very rarely look through such bodies, but only at them. Perhaps the commonest example where conditions both of regularity of structure of the screen and smallness of the source of light are met is that of an electric light seen through a silk umbrella. The type of figure here presented is that of an artificial star seen through three holes at the vertices of a right-angled triangle.

But there is a very large class of examples of such color effects produced in a different manner yet with precisely the same theoretical explanation. We have seen, on page 9, that optically a small mirror is exactly the same as a hole of its own size through which we can see a perverted image of all that lies in front of its plane. Consequently, looking toward a number of small mirrors is optically the same as looking *through* a number of like apertures, except that the source or sources of light appear to be in directions other than their real ones. In its action with respect to light reflected from it a piece of polished glass or metal with close, uniformly distributed grooves would resemble that of a series of fine slits in an opaque screen with respect to transmitted light. When held close to the eye a small source of light would appear flanked on either side by a series of colored images; at a distance from the eye it would appear of the color proper to the kind of waves that are not mutually self-destructive in the direction of the eye. This peculiar effect, which depends upon the direction in which the grooved reflector is observed and upon the direction of the source of light, is called iridescence. Mother-of-pearl, which is composed of smooth reflective layers, shows this property in a marked degree, on account of its peculiar structure. That it is due to the structure alone may be proved by taking a print of a piece of bright mother-of-pearl on white wax, when it will be found that the surface of the wax also shows irides-

cent colors. Feathers possessing at the same time regular structure and brilliant lustre exhibit the same phenomenon. Most of the beauty of peacock's feathers, and all of that of the iridescent feathers of the male birds of the turkey, pigeon, and humming-bird families are thus explained. If the source of light is very large or if the structure of the reflecting surface is somewhat irregular, the colors are less pronounced or may be wholly wanting, the only effect being a remarkable change of lustre with varying obliquity. The sheen of white mother-of-pearl, of the gem called cat's-eye, as well as that of satin spar, is of this kind. The variety of feldspar known as labradorite has such regularity of fibrous structure that a polished surface illuminated by a source of light of moderate extent will show most vivid colors.

There is a difference in the arrangement of colors produced by a prism and by interference, which is characteristic, and worthy of note. In the former case the blue end of the spectrum is further away from the undeviated position of the light than the red, while in the latter case this rule is reversed. This difference may enable one to distinguish in some cases the origin of a color. For example, the corona seen about the moon when covered by a very light cloud is blue within and red on the edge more remote from the moon; on the other hand, the large colored circle which is called a halo, and which is not infrequently seen around the sun or moon in cold weather, has a red border on the inner edge and a blue one on the outer. According to the foregoing principle, we ought to conclude that the first is due to interference and the second to refraction. This is really the case, for the corona is produced by the diffractive action of small and uniform drops of water (which act in the same way as a number of holes of the same diameter), and the halo is due to the refraction of the light by small sixty-degree prisms of ice suspended in the atmosphere.

There is another method of securing interference of light which, though it has no bearing on the theory of vision or

on the theories of optical instruments, is of such common occurrence that we are justified in entering upon its treatment here. Instead of using two portions of the same wave-surface, we may divide a system of waves into two separated systems and allow them, after having traversed different paths, to reunite. There are numerous ways of doing this, but we shall confine our attention for the present to the most common method.

Suppose a system of light waves to fall upon a very thin plate of some transparent substance, water, for example; the waves would be reflected in part from the first surface, but the greater portion would pass on through the plate until it fell on the second surface, where another part would suffer reflection, and, with slight loss at the emergent surface, re-enter the air. This second system of waves would be obviously behind the first by twice the thickness of the plate, if the reflection were nearly normal. Suppose now that the plate has a thickness of one-fourth of a wavelength of red light, then the two systems would be mutually destructive as regards these particular waves, so that their effect on the retina would be that of white light minus red, or, as we shall see in the second part of this book, it would appear green-blue. If the plate were a little thinner, say one-fourth the yellow wavelength, this color would be absent in the combined effect, and the plate would appear blue. Again, suppose the plate to have a thickness of three-fourths of a wavelength of red light, then the retardation of the waves reflected from the back surface would be three-halves of a wavelength, and mutual destruction of this particular wavelength would ensue, with a corresponding modification of the color. This is essentially the explanation of the production of colors by reflection from thin plates, such as soap-films or thin layers of oil on the surface of water. To make the explanation quite in accordance with fact, we should take into account the curious change in phase that waves undergo when reflected at the passage from a rarer to a denser medium, which does not occur in reflection under reversed condi-

tions. But as this fact does not cause any change ordinarily perceptible, we may ignore it here.

There is one notable peculiarity of the colors produced in this way which is of great interest, not only because it affects the whole character of the phenomena, but because it renders evident at once why we see these colors in very thin plates alone. It depends upon the fact that the colors are not those of the various wavelengths, as in the case of the prismatic spectrum, or even those of the sum of several wavelengths, as is true of many of the interference phenomena last considered, but they are the colors proper to white light after being deprived of one or more definite wavelengths. Let us suppose that we have a plate of such thickness that the yellow is eliminated in the reflected light; the remainder is blue, since these two hues are complementary. But if the thickness is such that the retardation of the system of waves from the second reflection is three-halves of a wavelength of red light, it will be five-halves of a wavelength of the shorter green waves; hence the color will be white minus such red and green. Now red and green light combined form yellow; therefore white deprived of these two hues will be blue also, only the blue is notably paler than that from the thinner plate. These two resultant colors are called blues of the first and of the second order, respectively. An extension of the reasoning shows that a still thicker plate may eliminate three wavelengths, which would yield again a still paler blue, provided that the three combined would produce a yellow hue. Should the increase of thickness of the plate be so far extended that certain wavelengths of all colors are eliminated, that is, that a series of colors which combined would themselves produce white, the sum of the remaining colors would be also white. Thus we see that only films of a very few wavelengths in thickness can produce colors by reflection, and also that there is a limit to the number of orders of the recurrent colors.

All these consequences of theory may be observed and verified very conveniently by soap-films, either by watching a

moderate-sized soap-bubble protected from currents of air, or, better still, by dipping a wire ring two or three inches in diameter into a solution of soap, whence it can be removed with a film across it. If this is held in an inclined position where the reflection of the sky from it can be seen, a series of colored horizontal bands will appear, of which the upper ones correspond to the thinner plate. It will be observed that the hues repeat themselves, but grow paler with each repetition. Sir Isaac Newton, who first studied the phenomena of thin plates and their causes, devised an apparatus for showing all the colors with perfect regularity. It consisted of a polished spherical glass surface of very small curvature, pressed against a flat polished plate. When skylight is reflected from the thin plate of air included between these two bodies, a series of concentric rings may be seen, of which the inner ones are those corresponding to the thinner portion and possess the more intense coloring. These rings are universally known as Newton's rings.

Since the colors produced in the manner described are to be seen by looking directly at the reflecting plate, not through or beyond as in the cases of interference previously considered, they are much more frequently observed. Besides the cases already mentioned, they may be often seen in partly fractured glass or ice. Quartz and other crystals not uncommonly have internal fractures which show brilliant colors of this origin, and the magnificent play of colors in the precious opal is explained in the same way. The surface of glass which has been long exposed to dampness undergoes a change which gives the effect of a thin layer of transparent refracting substance superimposed upon it. In the famous antique glasses of Cyprus the layers seem to be multiple and the effects are wonderfully brilliant. This is imitated with a certain amount of success in modern glass, by exposing the material to corrosive vapors at high temperatures.

If a piece of polished steel is exposed to the air when heated, the surface is oxidized to a varying degree of depth, depending primarily on the temperature. This thin sheet of

oxide reflects light as a thin plate, and affords a valuable aid to the mechanic in judging the proper temperature for tempering. A layer of such thickness as to eliminate the green waves, thus leaving a brilliant purple, is often used as an ornamental finish for small steel work, such as screws, the hands of a watch, and so forth.

CHAPTER V

THE TELESCOPE

There is no instrument which has done so much to widen the scope of human knowledge, to extend our notions of the universe, and to stimulate intellectual activity, as has the telescope. Certainly no other instrument except the microscope could for a moment be held to dispute this preeminence, but the former instrument is incomparably more interesting in its history in the same degree that its history is more simple and more comprehensible. To trace its development from a curious toy in the hands of its discoverer to the middle of this century, is to be brought into contact with most of the great philosophers who have achieved eminence in physical science, from the time of the Renaissance. Galileo, Torricelli, Huyghens, Cassini, Keppler, Newton, Euler, Clairault, the Herschels, father and son, Fraunhofer, Gauss — these form only a portion of the list of great names. The growth of the telescope toward perfection has constantly carried with it increase of precision in astronomy, in navigation, and in all branches of engineering. It would be easy to show that even pure mathematics would be in a far less forward state had there been no problems of astronomy and physics which were first suggested by the employment of this instrument. It is to a review of its history that this chapter is to be devoted. In it we may hope to find a succinct statement of the origin and development of this potent aid in the study of nature, a mention of some of the more important achievements depending upon it, and a sketch of its gradual improvement to the magnificent and complicated instrument which constitutes the modern

equatorial. Following this we shall strive to gain an idea of the imperfections which the generations of ingenious artisans have had to contend with in attaining its present degree of perfection, and possibly to forecast in a manner what the future may bring to us in the same field.

In Chapter II. was sketched the early history of the telescope, immediately after its discovery by Lippershey and its application to scientific investigation by Galileo. That this discovery was really an accident we may be quite sure, for not only was there no developed theory of optics at that time, but even the law of refraction, which lies at the basis of such theory, was wholly unknown. So also it seems more than probable that Galileo's invention was empirical and guided by somewhat precise information originating in Holland, such as that the instrument consisted essentially of two lenses of which one was a magnifying and the other a diminishing lens; at least, that Galileo's telescope was like that of Lippershey; that, theroretically considered, it is not as simple as one made of two magnifying lenses, as is evinced by the fact that the first philosopher to establish an approximate theory of optical instruments invented the latter and prevailing form; and finally, that Galileo published no contributions to optics — all these reasons together seem quite sufficient to enforce such a belief. But in any case Galileo's merit is in no wise lessened by having failed to do what could not have been done at that time; and to him more than to any of his contemporaries clearly belongs the glory of having made the telescope the most efficient servant of scientific research which human ingenuity has yet invented.

No further discoveries of great moment were made until over a generation after Galileo proved the existence of spots on the sun, in 1611. This cessation of activity was doubtless owing to the difficulty in securing telescopes of greater power than those possessed by Galileo, which he would hardly have left until their capacity for discovery had been fully exhausted in his own hands. By the middle of the seventeenth century, however, several makers of lenses had

so far improved the methods of grinding and polishing glass that telescopes notably superior in efficiency to the best of Galileo's were procurable. Of these Torricelli, Divini, and Campani; the French Auzout, who constructed a telescope 600 feet long, although no means was ever devised for directing such an enormous instrument toward the heavens; and above all, Huyghens, have won distinction as telescope-makers. The last-named philosopher, by means of a telescope of his own construction, discovered, in 1655, the largest satellite of Saturn, thus adding a fifth member to the list of planetary bodies unknown to the ancients. But his most important astronomical discovery, made also in 1655, was the nature of the rings of Saturn. This object had greatly puzzled Galileo, to whose small telescope the planet appeared to consist of a larger sphere flanked on either side by a smaller one; but when, in the course of the orbital motion of Saturn, the rings entirely disappeared, he was wholly unable to suggest any explanation for so unexpected a phenomenon. The planet had thus presented a remarkable problem to all astronomical observers for more than forty years, and the records of the efforts to solve it during that interval afford us a most excellent means of estimating the progress in practical optics. Huyghens announced these discoveries in 1656, but that relating to the ring was given in the form of an anagram, the solution of which was first published in 1659. This latter discovery was contested in Italy by Davini, but was finally confirmed by members of the Florentine Academy with one of Davini's own telescopes.

A few years later the famous astronomer Cassini, having come to Paris from Italy as Royal Astronomer, commenced a series of brilliant discoveries with telescopes made by Campani of Rome. With these, varying in length from 35 to 136 feet, he discovered four satellites of Saturn in addition to the one discovered by Huyghens. The whole number was increased by Herschel's discovery of two smaller ones in 1789, a hundred and five years after Cassini's last discovery, and again by Bond's detection of an eighth in 1848. The

Saturnian system, to which the telescope has doubtless been directed more frequently than to anything else, thus serves as a record of the successive improvements of the telescope. Highly significant is the fact that the discoveries of the eighteenth century were made with a reflecting telescope, the others all being with refracting instruments.

Cassini's discovery of the two satellites now known as Dione and Tethys, in 1684, was not accepted as conclusive until long afterward, when Pond, in 1718, with a telescope 123 feet long, which Huyghens had made and presented to the Royal Society, saw all five of the satellites. This particular instrument is of especial interest, because it is the only one of those of the last half of the seventeenth century which has been carefully compared with modern instruments. Moreover, it is without doubt quite equal in merit to the best of that period. But we find that, although it had a diameter of six inches, its performance was hardly better than that of a good modern instrument of four inches in diameter and perhaps four and a half feet in length, while in regard to convenience in use the modern compact telescope is incomparably superior.

Another notable discovery of this period was that of the duplicity of the rings of Saturn, made by Cassini in 1675. It is often said, apparently after a statement in Grant's History of Astronomy, that this discovery was made a decade earlier by the brothers Ball in England; but it is quite impossible to find authority for this conclusion from the original descriptions of their observations in the Transactions of the Royal Society where they were published. Of more interest to us is the fact that these observations appear to have been made with a telescope 38 feet long, of English manufacture; if this is true, it seems to be the first notice of the existence of opticians of high merit in that country. We must regard Cassini's discovery of the third and fourth satellites of Saturn, however, as marking the very furthest reach of the old form of telescope; a century was to elapse and an entirely new form was to be developed before another consid-

erable addition to our knowledge of the aspect of the heavenly bodies was to be made. It is true that larger telescopes were made, and Huyghens invented a means by which they could be used without tubes; but notwithstanding this improvement, they proved so cumbersome as to be impracticable.

The older opticians had found that if they attempted to increase the diameter of a telescope they were obliged to increase its length in a much more rapid ratio to retain the same clearness of vision. The reason for this was not clearly

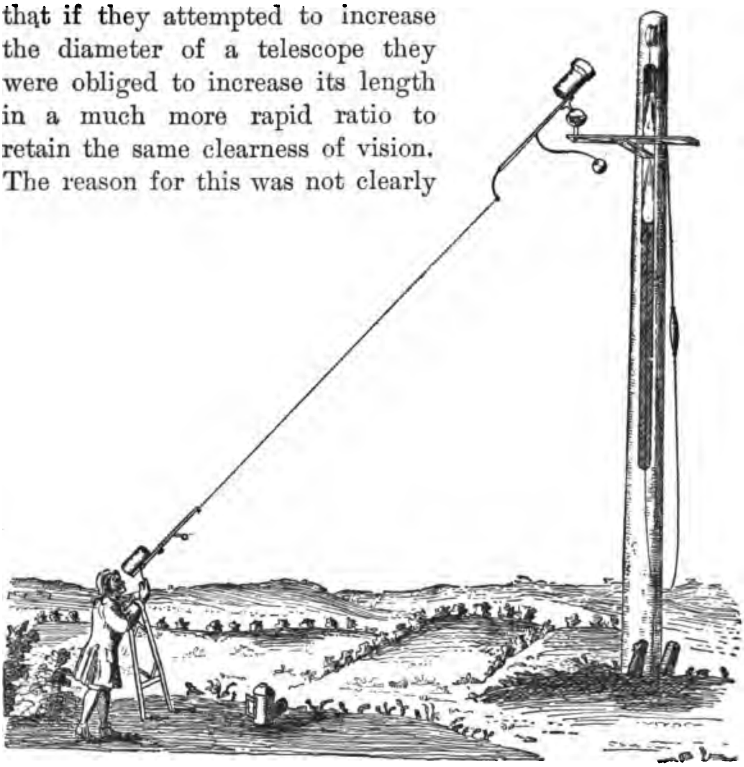

FIGURE 17.

understood, but it was supposed to be owing to the fact that a wave-front changed in curvature by passing through a spherical surface ceases to remain strictly spherical, as has been shown on page 15 of this book. This deviation in shape of the refracted wave from a true sphere is called spherical aberration. When the refracting surfaces are large

and of considerable curvature, this error becomes very serious; but by using small curvatures, which in a telescope obviously corresponds to great length, the effects of the error can be made insensible; hence the rapid increase of ratio of length to diameter remains unaccounted for. Newton's discovery of the composite nature of light and of the phenomenon of dispersion enabled him to explain the true cause of indistinctness in short telescopes, namely, that the refraction by the objective varies for different colors; hence, if the ocular is placed for one particular color it will not be in the right position for any others, whence the image of a star or planet will seem to be surrounded by a fringe of color. Newton found this source of indistinctness in the image, which is now known as chromatic aberration, many hundreds of times as serious as the spherical aberration. As he was persuaded by his experiments that this obstacle to the further improvement of the refracting telescope was insuperable, he turned his attention to a form of telescope which had been suggested a number of years earlier, and in which the image was to be formed by reflection from a concave mirror instead of by refraction. Since reflection is unattended by separation into elementary colors, it is evident that this method would eliminate the chief defect of the refractor. With his own hands Newton constructed a small telescope on this principle, and it still exists in the possession of the Royal Society. This little instrument seems to have been of about the same power as Galileo's telescope with which he discovered the satellites of Jupiter, but it was hardly more than six inches in length.

Since that time the reflecting telescope has had a remarkable history of development in the hands of a number of most skilful mechanicians, who have also for the most part been distinguished by their discoveries in physical astronomy. We may therefore advantageously depart from the chronological order of treatment, and follow the history of this type of instrument to the present time. This course is the more natural because we may probably regard the supremacy of

the reflector, undisputed a century ago, as passed forever. Only in the province of astronomical photography does it still find defenders.

Even after Newton's invention was made public, little was done toward the improvement of telescopes for half a century, until, in 1723, Hadley presented to the Royal Society a reflector of his own construction, which was found to equal in efficiency the Huyghens refractor of 123 feet in length. From this epoch we may date the beginning of the supremacy of reflectors. A few years later Short commenced his career as a practical optician, and for thirty years he remained unsurpassed in the excellence of his instruments. During this time many telescopes more powerful than the best of the previous century, and infinitely more convenient to use, had been made and scattered throughout Europe, but there was also a singular dearth of telescopic discovery. Perhaps men thought that the harvest had already been gathered; or perhaps we may find the explanation in the fact that the great cost of telescopes so restricted their use that the impulse to discovery by their means was confined to a very small class. We might well regard the latter cause as the more probable one, in view of the remarkable manner in which the standstill in this branch of science was finally followed by a brilliant period of discovery rivalled alone by that of Galileo.

William Herschel was born in 1738, in Hanover. In 1755 he left his native country, and, going to England, secured a position as organist in Octagon Chapel, Bath, where we find him in 1766. Here he became so profoundly interested in the views of the heavens which a borrowed telescope of moderate power yielded him, that he tried to purchase an instrument in London. The cost of a satisfactory one proving beyond his command, he determined to construct a telescope with his own hands. Thus he entered upon a course which was to reflect honor upon himself, his country, and his age, and which led him to add more to physical astronomy than any other one man has added before or since. With almost

inconceivable industry and perseverance he cast, ground, and polished more than four hundred mirrors for telescopes, varying in diameter from six to forty-eight inches. This in itself would imply a busy life for any artisan; but when we remember that all this was merely subsidiary to his main work of astronomical discovery, we cannot withhold our admiration.

Fortunately for science as well as for himself, he made, early in his career, a discovery of the very first importance, which attracted the attention of all Christendom. On the night of March 13, 1781, Herschel was examining certain small stars in the constellation of Gemini, with one of his telescopes of a little more than six inches in diameter, when he perceived one star that appeared "visibly larger than the rest." This proved to be a new planet now known as Uranus. In the following year the discovery led to his appointment as Astronomer to the King, George III., with a salary sufficient to enable him to devote his whole time to astronomy.

One of the fruits of this increased leisure was the construction of an instrument far more powerful than had been dreamed of by his predecessors, namely, a telescope four feet in diameter and forty feet in length. Commenced in 1785, Herschel dated its completion as August 28, 1789, when he discovered by its means a sixth satellite of Saturn, and, less than a month later, a seventh even closer to the planet and smaller than the sixth. We may fairly regard this achievement as marking the limit of progress in the reflecting telescope, for although at least one as large is now in use, and one even half as large again has been constructed, it is more than doubtful whether they were ever as perfect as Herschel's at its best.

There has been one improvement in the reflecting telescope since the time of Herschel which ought not to be left unnoted here, namely, that of replacing the heavy metal mirror by one of glass rendered even more highly reflective than the old mirrors by a thin coating of silver deposited by chemical means upon the polished surface of the glass. The great

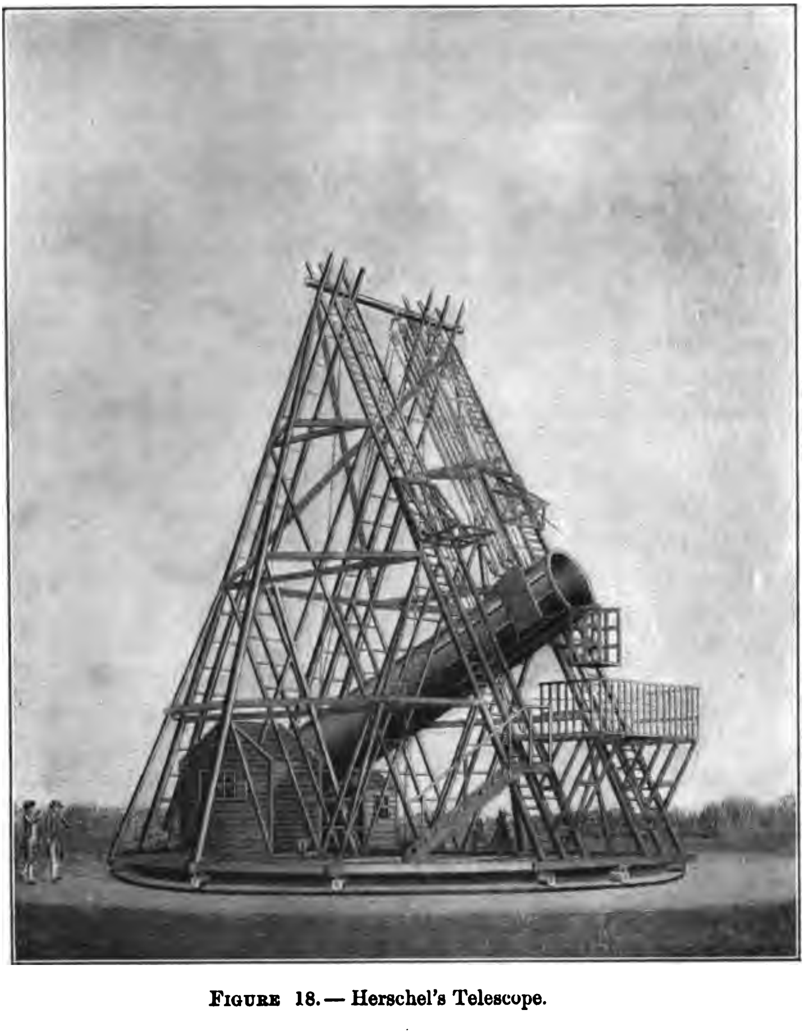

Figure 18.— Herschel's Telescope.

advantage of the modern form of reflector lies, not so much in the relative rigidity and lightness of the material, as in the fact that the surface when tarnished can be readily renewed by the simple process of replacing the old silver film by a new one; whereas in the metal reflectors a tarnished surface required a repetition of the most difficult and critical portion of the whole process of construction. Moreover, the making of such telescopes is so comparatively simple that an efficient reflector is far less expensive than is a refracting telescope of like power, so that this type may be regarded as particularly the amateur's instrument. On the other hand, such telescopes are, like their predecessors, extremely inconstant, and require much more careful attention to keep them in good working order. It is for such reasons, doubtless, that silver-on-glass reflectors have done so little for the advancement of astronomical discovery. In astronomical photography, however, they promise to do much on account of their absolute freedom from all color error and the ease with which large instruments can be made. In England some very remarkable photographs of nebulæ have been made in recent years with large reflectors, both by Mr. Common and by Mr. Roberts, the former using a reflector of three feet in diameter and the latter one of twenty inches. Although it is impossible to assert that no refracting instrument would yield as excellent results, it remains true that the work from these sources at the time of its publication proved a great surprise to all astronomers.

We now go back to a quarter of a century before Herschel discovered the new planet — to the very year, indeed, when that great astronomer first set foot on English soil — in order to trace the history of another form of telescope which has remained unrivalled for the last half-century in the more difficult fields of astronomical research, and which to-day finds its most perfect development in the instruments at Williams Bay, at Mt. Hamilton, at Pulkowa, at Vienna, and at Washington.

As a result from his experiments, Newton had declared

that separation of white light into its constituent colors was an inevitable accompaniment of deviation by refraction, and consequently the shortening of the unwieldy refractor of his time was impracticable. The correctness and validity of the experiments remained unquestioned for nearly a century, but finally a famous German mathematician, Euler, attempted to throw doubt on Newton's conclusion in a singular way. His argument was that, since the eye does produce colorless images of white objects, it might be possible by the proper selection of curves so to combine lenses of glass and of water as to produce a telescope free from the color defect. Although Euler's premise was an error, since the eye is not free from dispersion, his efforts had the effect of leading to a much more critical study of the phenomena involved. In this John Dolland, an English optician, met with brilliant success. Repeating an experiment of Newton's with a prism of water opposed by a prism of glass, he found that deviation of light could be produced without dispersion into prismatic colors. More than this, he found that the two varieties of glass then, as now, common in England — namely, crown or common window glass, and flint glass, which is characterized by the presence of a greater or less quantity of lead oxide — possessed very different powers in respect to dispersion. Thus, of two different prisms of these two varieties of glass which would deflect the light by the same angle, that made of flint glass might form a spectrum quite twice as long as that formed by the other. Hence, if a prism of crown glass deflecting a transmitted beam of light by ten degrees were combined with one of flint glass which would deflect the light five degrees in the opposite direction, there would remain a deflection of five degrees without division into colors. It also follows that a positive lens of crown glass combined with a negative lens of flint glass of half the power would yield a colorless image. Such combinations of two or more substances are called achromatic systems. Some of them are illustrated in Figure 23, page 97, where the shaded portions represent the flint glass.

It is a singular fact, which is worthy of note at this point, that more than twenty years before Dolland's success achromatic telescopes had been invented by Mr. Chester Moor Hall and constructed for him; but for some reason now difficult to explain, the possibility of such construction remained unknown to the world of science until after Dolland's telescopes had attained fame.

Why the greatest philosopher who ever lived, and at the same time one of the most successful experimenters, should have been corrected by a follower employing the same method of investigation, has never, to my knowledge, been explained. Still, the reason does not seem far to seek. At the time of Newton the two substances which stood at the extremes of refractive power, and which at the same time were so abundant as to be readily procurable for investigation, were water and common glass. There were, of course, a vast number of gems and other rare substances which stand quite outside these limits, but these were not suited for the methods of research at command at that time. Now it so happens that these two substances, water and common glass, do have essentially the same dispersive power, so that when a prism of glass is corrected by a prism of water in the reverse position, the deviation vanishes with the color. This being so, not only was it natural to conclude that the indicated deduction was a general law, but it would have been unphilosophical to doubt it without further experience. It is quite conceivable that had flint glass been as familiar a substance then as it was a century later, the true relation of dispersion to refractive power, or, rather, their independence, would not have escaped the earliest and most acute investigator in this field.

For a long time this ingenious invention of Dolland remained fruitless for astronomical discovery, although it was early applied to use in meridian instruments. This was due to difficulty in securing sufficiently large and perfect pieces of glass, more particularly of flint glass, to meet the demands of the optician. Not until after the beginning of

the last century was any real advance exhibited in this branch of the arts. Even then success appeared, not in England or France, where strenuous efforts had been made to improve the quality of optical glass, but in Switzerland. There a humble mechanic, a watchmaker named Guinand, spent many years in efforts, long unfruitful, to secure a method of making large pieces of optically useful glass. Just what degree of success he attained there we do not know, though from the fact that, during this period of twenty years of experimenting, good achromatic telescopes of more than five inches in diameter were unknown, we must conclude that his success was limited. In 1805 he joined the optical establishment of Fraunhofer and Utschneider in Munich. Here he remained nine years, and with the increased means at his disposal and with the aid of Fraunhofer he perfected his methods so far that the production of large disks of homogeneous glass became only a matter of time and cost. It is not too much to say that all the large pieces of optical glass which have since been produced and which have added so incalculably to the resources of the scientific investigator, whether made in Germany, France, or England, owe their existence to the wonderful patience and industry of this Swiss watchmaker, whose trials and achievements recall those of Palissy.

Fraunhofer was a genius of high order. Although he died at the early age of thirty-nine, he had not only brought the achromatic telescope to a degree of optical perfection which made it a rival of the most powerful of the reflectors, and so far improved upon the method of mounting it that his system has displaced all others, but he also made some capital discoveries in the domain of physical optics. His great achievement as an optician was the construction of an achromatic telescope nine and one-half inches in diameter, with which the elder Struve made his series of remarkable discoveries and measurements of double stars at Dorpat. The character of Struve's work demonstrates the excellence of the telescope, and shows us that it should be ranked as the equal of all but the very best of its predecessors. Indeed it

may be fairly concluded that not more than one or two telescopes of greater power, and those made and used by Herschel, had ever existed, while in convenience for use the new refractor was vastly superior.

For a long time Fraunhofer and his successors, Merz and Mahler, from whom the second equatorial of Pulkowa and the fifteen-inch telescope of the Harvard Observatory were procured, remained unrivalled in this field of optics. But they have been followed by a number of skilful makers whose products, since the middle of the century, have been scattered throughout the world. In Germany, Steinheil and Schroeder; in France, Cauchois, Martini, and the Henry brothers; in England, Cook and Grub; and in this country, the Clarks and Brashear — each has produced one or more great telescopes which have rendered his name familiar to all readers of astronomical literature. Of these the Clarks, father and son, have beyond a doubt won the first place, whether determined by the character of the discoveries made with their instruments or by the fact that the two most powerful telescopes in existence — the refractor of three-feet aperture of the Lick Observatory in California, and the forty-inch telescope of the Yerkes Observatory in Wisconsin — were made by them. The most notable discoveries made with their telescopes are the satellites of Mars, the fifth satellite of Jupiter, and the companion star to Sirius; but besides these there is a long list of double stars of the most difficult character discovered by the makers themselves, by Dawes in England, and, far the most notable of all, by Burnham in our own country.

The Yerkes telescope, which now stands as the highest representative of an art itself the growth of nearly three centuries of human endeavor, is represented in the accompanying picture. It is vastly more complicated than its famous Fraunhofer prototype, but it does not differ in any fundamental principle. Like that, it has an axis parallel to the axis of rotation of the earth, to the upper end of which there is attached another axis at right angles to it; the latter

FIGURE 19. — Fraunhofer's Equatorial Telescope.

carries at one end the telescope tube, which is supported not far from its middle point. If the instrument is rotated on the first described axis, which is called the polar axis, its line of sight will describe a circle on the heavens parallel to the equator; it is this property which gives the name equatorial to such a mounting. Since, in their diurnal motion, the heavenly bodies appear to describe circles parallel to the equator, to follow a star once found in the telescope it is only necessary to rotate the instrument on the polar axis alone and at the same rate as the apparent motion of the star. This motion can be secured with wonderful regularity and smoothness by a small motor — shown in part within the hollow cast-iron pier — called a driving clock. To each axis is attached a large divided circle which enables an astronomer to find any object in the heavens whose place is known, or to determine the position of a previously unknown object when brought to the centre of the field of the telescope. Important accessories to a large instrument are one or more small telescopes attached to the eye end, with their lines of sight parallel to that of the great telescope. These enable the astronomer to find an object and point his large instrument with greater facility, and are appropriately called finders.

A comparison of Figures 17 to 20, inclusive, is highly instructive. In the first we have the telescope of the seventeenth century in the state of highest development. Herschel's great telescope was not only the most powerful of the following century, but it was also perfectly typical of the best that that century produced. Finally, the third and fourth of these figures represent, respectively, the earliest and the most recent of modern powerful refractors, separated in time of construction by almost exactly three-fourths of a century.

There are two points in the theory of the telescope which are of fundamental importance, and which, by means of the theoretical considerations developed in the preceding chapters, can be readily understood. The first determines the range of magnifying powers for any given instrument, and

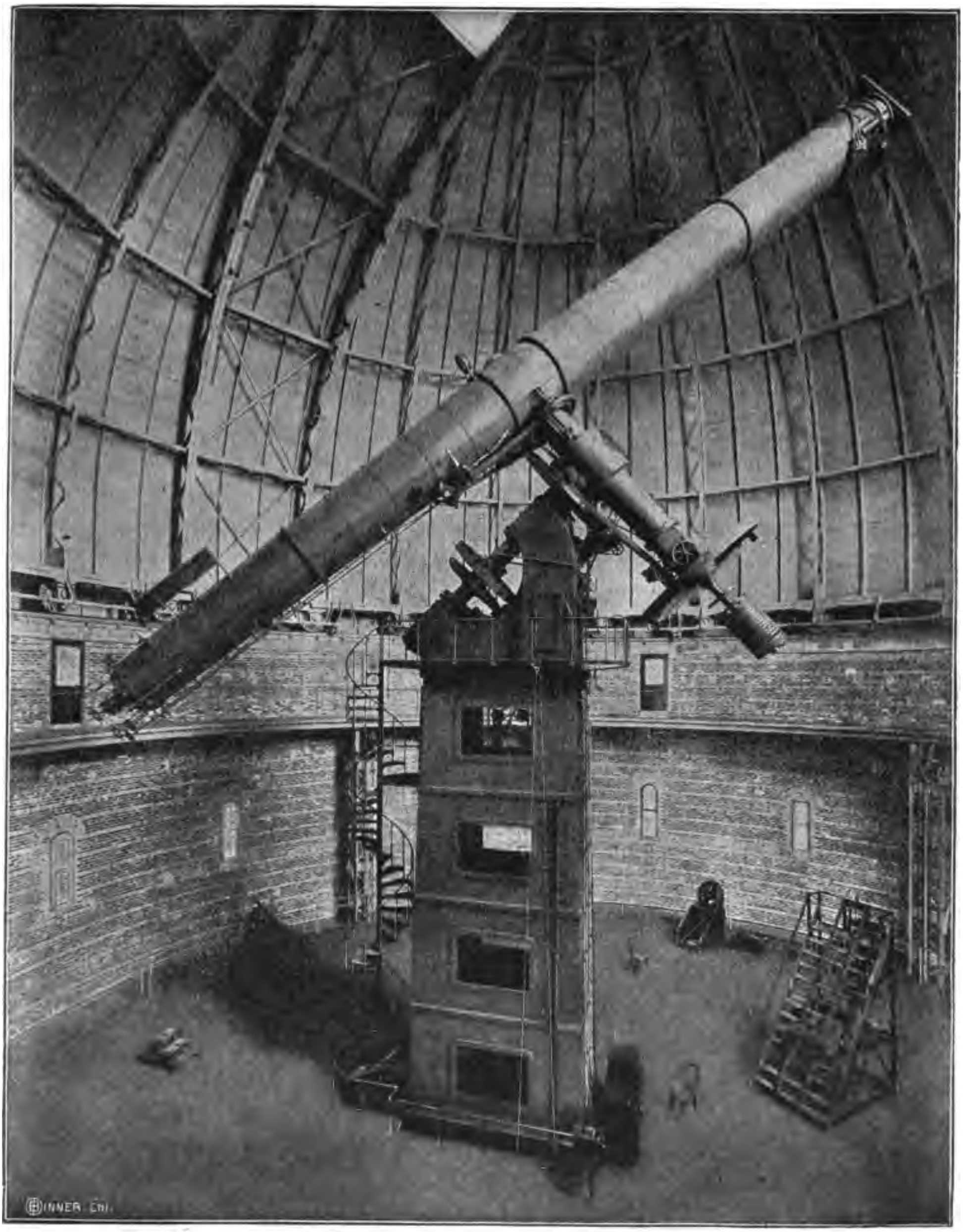

Figure 20. — Equatorial Telescope of the Yerkes Observatory.

the second enables us to find the absolute efficiency of a telescope.

In Chapter II. the essentials of a telescope were shown to be an object-glass of low power for forming an image, and an eyepiece, or ocular, for magnifying this image. But this ocular will form a real image of everything in front of it which is sufficiently remote; it will consequently form an image of the object-glass. Such an image can be always seen close to the ocular if the telescope is directed toward a bright surface, as, for example, toward the sky. Now a simple calculation shows that the ratio of the diameter of the objective to that of this ocular image is exactly equal to the magnifying power of the telescope. This is by far the best method of determining the magnification. This ocular circle is also the smallest area through which the light passes after leaving the telescope, and thus marks the best position of the eye of the observer. If the pupil of the eye is smaller than the ocular circle, it is obvious that not all the light which is transmitted through the objective will reach the retina. The lowest power, therefore, which can be used with full advantage on any telescope is found by multiplying the diameter of the objective in inches by five or six. Thus the lowest power which can be used with the Lick telescope and have it retain its superiority over smaller instruments is from 180 to 220.

On the other hand, experience teaches that vision begins to be sensibly impaired when the ocular circle is much less than one-thirtieth of an inch in diameter; consequently the highest power which is generally available is found by multiplying the diameter of the objective in inches by thirty or forty. For example, the magnifying powers employed on the twenty-six inch equatorial at Washington are practically confined to 400, 600, and 900, making the diameters of the corresponding ocular circles 1/15, 1/23, and 1/35 of an inch. It was with the lowest of these powers that both the satellites of Mars were discovered by Professor Hall.

Thus we see that the range of useful powers in any

telescope is practically confined to from five to forty times the number expressing the diameter of the objective in inches. The lowest limit has been previously established by theoretical considerations and does not admit of question; but, as far as appears in what precedes, the upper limit is quite empirical, and naturally suggests the question as to whether we may not in the future so perfect the telescope that smaller instruments can bear higher powers and thus do the work of greater ones. The answer to this eminently practical question is an unqualified negative. The following reasoning demonstrates that the upper limit of magnification is fixed by the nature of light itself, and increase of power necessarily involves an increase of size in the telescope.

It is obvious that a telescope directed to a star gives precisely the condition represented in Figure 14, page 32, that is, a concave wave-front with its centre at the place which marks the image of the star. The diameter of the wave-front just after passing the objective, represented by the distance *ab* in that diagram, is the same as that of the objective. Hence, it follows that the image of the star must consist of a disk surrounded by a series of concentric rings precisely similar in appearance to the artificial star seen through a needle hole in a card. In the case of a slit of width *ab* it was there shown that the distance from the centre of the image to the nearest dark space is equal to *ap*, the radius of the wave-front, multiplied by the length of a wave of light and divided by *ab*. The more intricate analysis for a circular aperture shows that for this case the term should be multiplied by 1.2. Thus the image of a star in a telescope consists of a disk whose *diameter* is 2.4 $F\lambda/D$, where F is the focal length and D the diameter of the objective. The angular diameter of the image is this quantity divided by F, which for a one-inch telescope, assuming that the length of a wave of light is 1/46000 of an inch, would be 10″.75; hence two stars separated by an interval of 10″.75 would just appear to touch, if the light could be traced quite to the place of the dark ring. In fact, however, the edge

of such a disk is so faint that it appears much smaller than indicated by this reasoning, and, under the most favorable circumstances, at somewhat less than half this distance two stars can be just seen as two distinct objects. Consequently, to find the closest double stars which can be seen with any perfect telescope we divide $4''.56$[1] by the diameter of the objective expressed in inches. This rule is found to be quite accurate for even the largest telescopes, if perfect; for example, under favorable circumstances, the Washington telescope of twenty-six inches diameter ought, according to the rule, to divide stars separated by an angular interval of $0''.20$ only, and in fact double stars $0''.23$ apart have been seen as separated.

It is clear that if the images of two points are not separated they cannot be made to appear so by mere increase of magnification by the ocular, which is quite analogous to the familiar fact that beyond a very limited extent one can see nothing more in a photograph by increasing the magnification under which it is examined. The angular separation is thus shown to depend on the diameter of the objective alone, and it is this element, therefore, which determines the upper limit of the power of a telescope as well as the lower limit.

The preceding pages make quite evident that the present degree of excellence in the telescope has been attained only after prolonged efforts, extending through a period of many generations and participated in by a host of inventors, among whom are a number who have been illustrious in other fields of work. It will not be without interest to consider the reasons why this progress has been so slow in appearance; why even now the greater number of telescopes in existence are distinctly bad when judged by a high standard; how it was that for a time, very brief it is true, the glassmaker was in advance of the demands of the optician; and, finally, why the first of the great modern objectives was in

[1] This is the value fixed upon by the astronomer Dawes after long experience with excellent telescopes.

the hands of the most skilful telescope-maker of Great Britain for seven years without satisfying that optician as to its ultimate excellence. These questions cannot be answered in a word, nor do they admit of complete answers; but much may be gained by recognizing that they involve two distinct kinds of reasons, namely, those depending upon purely technical difficulties of construction, and difficulties which, since they depend upon the solution of a definite mathematical problem, may be called theoretical difficulties. We shall consider these in turn.

The art of lens making can certainly be traced back to the thirteenth century, although the methods employed even as late as the beginning of the seventeenth were so crude that Galileo found the utmost difficulty in making a lens sufficiently good to bear a magnifying power of thirty times. We may be quite sure that he could command the aid of the best artisans in Europe and a complete knowledge of their methods; still he was obliged to train his own hands to attain a necessary and higher degree of skill than was possessed by any of his contemporaries. A most illuminating fact with regard to the relative state of the industrial arts in those times as well as in the present is that for a few cents it is now easy to secure a pair of spectacle lenses which would form a telescope rivalling if not surpassing in power that first and most famous of all telescopes. Not until the lapse of another generation was there such improvement in the technique of lens making that further advances in astronomical discovery became possible. This slow progress is to be explained by the extremely critical requirements for a good lens. A departure, by a fraction of a hundred-thousandth of an inch, of a material portion of the surface from the correct geometrical form will greatly impair the performance of an objective. But even at the present day the limit of accurate measurement may be set at about a one hundred-thousandth part of an inch, while it is quite probable that ten times that amount was vanishingly small to the artisan of a century and more ago. It was necessary, therefore, to devise a method

of polishing — for it was comparatively easy to grind a surface accurately — which should keep the surface true within a limit far transcending the range of possible measurements. Huyghens seems first to have succeeded in this by polishing upon a paste which was formed to the glass and then dried, and by employing only a portion of the centre of a large lens. In Italy, at about the same time, Campani developed a system that he most jealously guarded as a secret up to his death, after which it became known from an inspection of the tools and unfinished work in his shop. The peculiarity of his method consisted in polishing his lenses with a dry powder on paper cemented to the tool upon which the lens had been ground. This practice still survives in Paris to the exclusion of almost all others, and it is in many respects the best for work which does not demand the highest scientific accuracy. This qualification does not mean to imply that the Parisian opticians have failed to produce work of a high degree of excellence, but only that the method is not free from serious mechanical objections, and that nowhere else has it been employed by opticians who have attained distinction as telescope-makers.

Newton seems to have been the first one to make use of a method which has since been developed to a state of surprising delicacy. Casting about to find a means which should be sufficiently "tender," to borrow his own expression, for polishing the soft speculum metal of his reflecting telescope, he fixed upon pitch, shaped to the mirror while warm, as a bed to hold the polishing powder. But the enormous value of this and similar substances lies not so much in the comparative immunity which it gives from scratching, but to the fact that under slowly changing forces it behaves like a liquid, while under those of short duration it acts as though it were a hard and brittle solid. Thus it is possible with such a polisher slowly to alter the shape of a lens while on the tool, and to correct errors which are found by inspection of the optical image where they may be very sensible, although due to insensible departures of the lens surface from

the proper form. In view of the thinness of the layer which he used, it is perhaps doubtful whether Newton recognized this particular property of his pitch polisher, and it is also a question whether he contemplated the refinement of shaping his mirror to the form of a paraboloid of revolution, which was clearly indicated as the proper form and which necessarily demanded this property. However this may be, there is no doubt that the merit of perfecting the process and bringing it to its present condition belongs to the English of the eighteenth century and the early part of the last century. In the Philosophical Transactions of the Royal Society we find many long papers relating to this art, contributed by skilful and successful amateurs, and we may regard the technique of the art of lens making as essentially complete by the middle of the century just ended and as common property, so that success no longer depends upon the possession of some secret of manipulation.

In this hasty review of the development of the telescope, no one interested in the history of scientific progress can fail to be struck with the prominence of the three greatest philosophers of the seventeenth century. That each one of these men was content to employ his own hands in advancing a purely mechanical art, adds to our admiration rather than detracts from it. The interest is heightened by the knowledge that specimens of the handiwork of all three still survive, and are treasured possessions of various scientific societies. Who would not be eager to add to such treasures a lens made by Spinoza? It is hardly probable that this philosopher advanced an art which enabled him to secure a modest living while engaged upon those philosophical speculations which will remain a monument to philosophy as long as interest in intellectual affairs survives, but such a memento might possibly command even a more wide-spread interest than any of those named. We may also recall here, quite appropriately, the name of another great philosopher, a contemporary of both Newton and Spinoza, Descartes, who first published the true law of refraction, and proposed methods of

shaping lenses to forms which his high mathematical powers led him to think superior to the spherical form; but as his contributions have not proved of real value in the development of the art, we may be content with this brief mention.

The second type of difficulties, namely, those which have been styled theoretical, we can treat, without the language of mathematics, only in a very general manner; but even with this limitation we may hope to derive some idea of what the problems have been and, still more important from our standpoint, what we may hope for in future improvements.

The obvious requirement in a perfect objective is that light coming from a point in the object should be concentrated at a point in the image; but this requirement combined with a prescribed focal length may be defined by three conditions, namely, freedom from color error, absence of spherical aberration, and assigned optical power. Now consider what provisions are at command for satisfying these conditions. They are four surfaces which must be very nearly spherical but may have any radii we please, the two thicknesses, and the single distance separating the two lenses. These number seven, and exhaust the list for a binary system. As a matter of fact, however, on account of the cost of the material and the want of perfect transparency in glass, it is necessary to make large lenses as thin as possible; and even the element of separation of the two lenses is found to be unavailable in practice, for reasons which will be touched upon later; at least this statement is true for two-lens systems, which engage our attention at this moment. Thus there are left only the four radii which may be regarded as constants susceptible of arbitrary choice. But these are more than enough to meet three conditions, whence our problem in its given form is indeterminate, and it becomes necessary to add another condition for the sake of definiteness. Of course this new condition will be chosen so as to add some other desirable property to the objective, and it is in the question as to what additional property is most desirable that we find room for wide diversity of opinion. Clairault proposed to make the fourth

condition that the two adjacent surfaces should fit together and thus admit of cementing with a transparent cement. This construction is of great value in small objectives and is very largely used, but in large telescopes it is found impossible to cement the components of the objective without changing their shapes to such an extent as quite to spoil their performance. At a period still early in the history of the modern telescope (1821), Sir John Herschel published a very important paper, in which he made the fourth condition that the spherical aberration should vanish for objects at a very great distance and also for those at a moderate distance. In this paper he computed a table, afterward greatly extended by Professor Baden Powell, for the avowed purpose of aiding the practical optician. It was doubtless this feature that for a considerable time brought his construction, which is certainly not recommended by the formally stated condition, into more general use than any other. But as all these tables were derived from calculations which wholly disregarded the thickness and separation of the lenses, they could yield a rather rough approximation only, and it is quite possible that the discredit with which opticians have received the dicta of mathematicians concerning their instruments may have been in part due to this very fact. Fraunhofer, whose remarkable achievements in this field have already been touched upon, is said to have imposed the condition that the magnifications for all zones of the objective should be alike, and thus made the problem determinate; he does not seem to have anywhere published his method or results. This condition is of high importance where a considerable field of view is required, as in photographic telescopes, those used for transit instruments, and, in general, those designed for the better class of surveyors' instruments. It is a singular fact that in its ultimate solution this very philosophical condition corresponds to a second approximation with that of Herschel, which hardly pretended to be more than a convenient way of avoiding a difficulty in the problem due to indefiniteness; so true is this that no one could tell from inspection of a finished

objective whether one or the other condition were the guiding principle in the construction.

A particularly interesting proposal for the necessary fourth condition is that of Gauss; but to describe it will demand an immediate statement of a fact which is to receive attention a little later. We have already tacitly assumed that freedom from color error is always attainable; but this is so far from true that as a matter of fact only a very moderate approximation to this ideal state can be reached. A general statement of this kind may be made: In any achromatic objective the focal point of a given wavelength is found to agree with that of some other one wavelength, but with one only. Now the Gaussian condition was that the spherical aberration for two such different wavelengths should vanish simultaneously. This seems to have been a *tour de force* of a mathematician, not a sober suggestion of an improvement in construction, for in point of fact the solution yields a form which is far from being a good one. It was generally believed that this condition could not be fulfilled; therefore Gauss, who seemed particularly fond of doing what all the rest of the world deemed impossible, straightway did it.

Such are the more famous proposals for giving definiteness to the foregoing problem; but they do not exhaust the list. Others have been suggested by the astronomers Litrow and Bohnenberger, but a detailed description of their solutions would add little of essential interest. Considering that none of the solutions have predominating merit, it is hardly surprising that the practical optician has followed the line of least resistance and adopted a form which costs him less labor than the others and is sensibly their equal. By making the crown lens equiconvex the trouble of making one pair of tools is saved, which probably is quite sufficient to explain the vogue of that construction. Of course this reason should have no weight with the astronomer who seeks the best instrument it is possible to make; therefore, as there seems to be nothing else to recommend the construction, it will probably be used little in the future.

The reason for so much futile work on the theory of the telescope objective is not far to seek. It had always been tacitly assumed that the condition of color correction, one of those which serves to define the values of the arbitrary constants, was readily determinable — in fact, one of the data of the problem — whereas it is the finding of just this datum which has offered peculiar difficulties. Fraunhofer brought all the resources at the command of his exceptional genius to bear upon this point and frankly failed, although in the effort he made a splendid discovery which has assured a permanence to his fame no less enduring than the history of science itself — the discovery of the dark, or Fraunhofer, lines in solar and stellar spectra. Gauss proposed the condition that the best objective must be that which produces the most perfect concentration of light about the place of the geometrical image of a point, just as the best rifle practice is that which produces the maximum concentration of hits about the centre of the target. That this principle is inadequate appears at once from the consideration that even if we take as much as a tenth of all the light from an object and divert it wholly from the field of vision, the telescope may still be practically perfect. All of Herschel's telescopes did much worse than this. But if we take the same proportion of light and concentrate it in the immediate vicinity of the image, the telescope will prove worthless. The true principle, briefly stated, is that the *weighted* mean position of the light should correspond with the place of the geometrical image, the weights being assigned according to the relative importance of the different wavelengths for the particular end in view. For example, if it is desired to affect the retina it is necessary to recognize that the yellow waves are far more effective than those of other colors, and for purposes of photography this pre-eminence is transferred to the blue or violet waves.

Perhaps the true difficulty with most of the theoretical discussions of this important problem is this: There is in them no recognition of the *relative weight* or importance of unavoidable errors. At the very outset of his task the opti-

cian is confronted by the fact that absolute elimination of color error in a binary system is impossible, because no two substances are known which have dispersive ratios independent of the color. He can, it is true, reduce the color error of the old single-lens type of telescope hundreds of times, and hence the length of the telescope tens of times; but the fact that he must stop at a point far short of perfection puts an end to still further shortening, and in considerable telescopes leaves the minimum ratio of length to diameter not far from 15 to 1. This restriction being recognized, we may revise our limiting conditions. They now become, first, fixed focal length; second, best attainable color correction; third, freedom from spherical aberration for a particular wavelength of light. This still leaves, as before, one of the necessary four lacking. What should determine the choice of the fourth condition? Surely there is only one rational guide. Consider the residual errors and, with proper regard to their relative importance, make the final condition such as to reduce these errors to the smallest possible values. But the only remaining errors for an image in the axis are secondary color error and spherical aberration for colors other than that for which it is eliminated, or, more scientifically stated, chromatic difference of spherical aberration. As to which of these is the more grave defect depends upon the use to which the objective is to be put. If it is a high-power microscope objective, it is certainly the second. If it is an objective to be used for photographing at a considerable angular distance from the axis, the question loses its physical significance, since the consideration of eccentric refraction is excluded. But if the objective is to be for a visual telescope, there is no question that the defect of secondary color error is indefinitely more serious. The fourth and determining condition must therefore be improvement of color correction to the last attainable limit. With this end in view the improvement may be made in part by a careful selection of material, for the available glasses are by no means of unvarying imperfection in this particular.

The question as to what further improvement may be hoped for in the future of this most important instrument of research is a natural and interesting one. The absolute power has been shown to depend upon the diameter of the objective alone, if errors of construction are perfectly eliminated; hence the question leads at once to the consideration of the probability of securing useful disks of glass much larger than those hitherto employed, for we may with little doubt assert that the ablest opticians are competent to shape lenses notably larger than any now in existence, to the necessary degree of precision. It is obvious that no limit can be set to the possible achievements of glass-makers in the future; but we can state positively that were there no limit to the size of sufficiently homogeneous glass at command, we should finally reach dimensions of which the necessary weight would produce such distortions that further increase would carry with it no increase in optical efficiency. Whether we have begun to approach this limit imposed by the finite rigidity of glass is a question which cannot be answered with positiveness, but some observations with the largest refractor now in existence may be interpreted to point in this direction.

There remains to consider what an increase in perfection of telescopes of ordinary power can yield. If it were possible to secure two varieties of glass which should possess this highly desired independence of dispersive ratio and color, we might eliminate the most serious error remaining in the telescope, provided always that the other physical properties of the materials were suitable. The highly scientific glass-makers of Jena have carried on experiments to this end for many years, but, it must be confessed, with rather discouraging results. It is true that several pairs of glasses have been made which materially reduce the defect, but either one of the glasses has proved to be perishable or too hygroscopic for use, or, in the most promising case, the refractive powers differ so little that the chromatic difference of spherical aberration for telescopes of moderate length constitutes a defect worse than that corrected. The employment of three kinds

of glass, although it admits of a fairly complete solution of the problem, does not seem to promise to be useful in large telescopes, on account of the much greater length enforced and the increased inconvenience of reflections from the greater number of surfaces. The possibility of accomplishing the same end in other ways must be left to another place for discussion. At present we may assert with confidence that the procedure to secure the best results is to have separate instruments for visual and for photographic work, each designed for its special end. There is no doubt, however, that much worth having may be gained by a careful study as to the best attainable form with the materials to be used — a precaution too often neglected.

CHAPTER VI

THE MICROSCOPE

THE history of the microscope is far less easily traced than that of the telescope. This is chiefly because its physical theory is so much more complicated than that of the latter instrument. So true is this that its development to a state of almost ultimate perfection preceded all adequate theory of its action. It is easy to point out certain definite discoveries or inventions in the course of development, which may serve for fixing the epochs from which the history could be most logically traced; but since in nearly every instance these discoveries have been made quite empirically and utilized is trade secrets, it is impossible to establish their date accurately, or even, in many cases, to determine to whom the credit of discovery belongs. We shall meet with the names of a few philosophers, however, who have done something toward the improvement of this important instrument, and who have given the results of their labors to the world by publication. Such are Lister, Goring, Amici, Wenham, and finally Abbe, who has not only brought the theory of the microscope to a satisfactory state of completion, but has also, in conjunction with the manufacturers Zeiss, made a number of most notable improvements which will form topics for consideration somewhat later. Besides these, there is a host of most skilful artisans who for half a century have carried on a vigorous rivalry in the attainment of increased excellence. Some of these are notable for more than the admirable products of their skill which have materially advanced our knowledge of nature, for to them are due for the most part those inno-

vations in the principles of construction which have given a new impulse to the art. Such are Amici in Italy; Chevalier, Oberhaeuser, and Hartnack in France; Andrew Ross in England; Spencer and Tolles in this country.

We shall give here a brief sketch of the development of the microscope in regard to its optical improvement, disregarding the hardly less interesting history of its mechanical construction and of its accessories.

The property of yielding a magnified image of an object, possessed by a lens-formed body, was without doubt known to the ancients. At least, we know that they had the necessary skill in grinding and polishing transparent gems, and there remain to this day many examples of antique cameos of such exquisite detail in finish as to force the conclusion that the artist must have been aided by a magnifying glass. Nor, considering the extreme meagreness of references to technical methods in mechanical arts by classical writers, is the absence of all contemporary allusions to such aids to be regarded as significant.[1] But however true the conclusion may be in regard to this knowledge of the ancients, we do know that not until after the middle of the seventeenth century did the microscope add greatly to our knowledge of nature. It is highly interesting that the marvellous scientific activity of this century opened two great realms for study — that of the boundlessly great and that of the invisibly small. Not less interesting is the wide difference between the rates at which exploration in these new fields was pushed. At the beginning of the nineteenth century the truths attainable as a result of mere increase of telescopic range of vision had been almost completely gathered, while those made accessible to us by the microscope were only fairly begun. The reason for this disparity in progress in the two fields will appear implicitly as we trace the history of the latter instrument.

[1] Natural lenses formed by hardened drops of gums and resins must have always been familiar objects, and amber was a substance frequently finished with curved surfaces in ancient times. It seems quite incredible that the magnifying properties of such bodies could have been unknown.

Almost the first discoveries of importance made with the microscope were published in a famous book by Robert Hooke, in 1665, entitled "Micrographia." The book is admirably illustrated by numerous engravings on copper and bears a quaint dedication to King Charles II., which suggests many reflections regarding the changes in the character of scientific publications since that time. The dedication reads as follows: —

TO THE KING

SIR, — I do here most humbly lay this small Present at Your Majesties Royal feet. And though it comes accompany'd with two disadvantages, the meanness of the Author, and of the Subject; yet in both I am incouraged by the greatness of your Mercy and your Knowledge. By the one I am taught, that you can forgive the most presumptuous Offendors: And by the other, that you will not esteem the least work of Nature, or Art, unworthy your Observation. Amidst the many felicities that have accompani'd your Majesties happy Restauration and Government, it is none of the least considerable, that Philosophy and Experimental Learning have prosper'd under your Royal Patronage. And as the calm prosperity of your Reign has given us the leisure to follow these Studies of quiet and retirement, so it is just, that the Fruits of them should, by way of acknowledgement, be return'd to your Majesty. There are, Sir, several other of your Subjects, of your Royal Society, now busie about Nobler matters: The Improvement of Manufactures and Agriculture, the Increase of Commerce, the Advantage of Navigation: In all which they are assisted by your Majesties Incouragement and Example. Amidst all those greater Designs, I here presume to bring in that which is more proportionable to the smallness of my Abilities, and to offer some of the least of all visible things, to that Mighty King, that has establisht an Empire over the best of all Invisible things of this World, the Minds of Men.

Blah!

Your Majesties most humble
and most obedient
Subject and Servant,
ROBERT HOOKE.

Hooke's observations were made chiefly with a compound microscope of his own design and construction, which is represented in the accompanying figure copied from his engraving. The lenses consisted of a small one at the lower end of

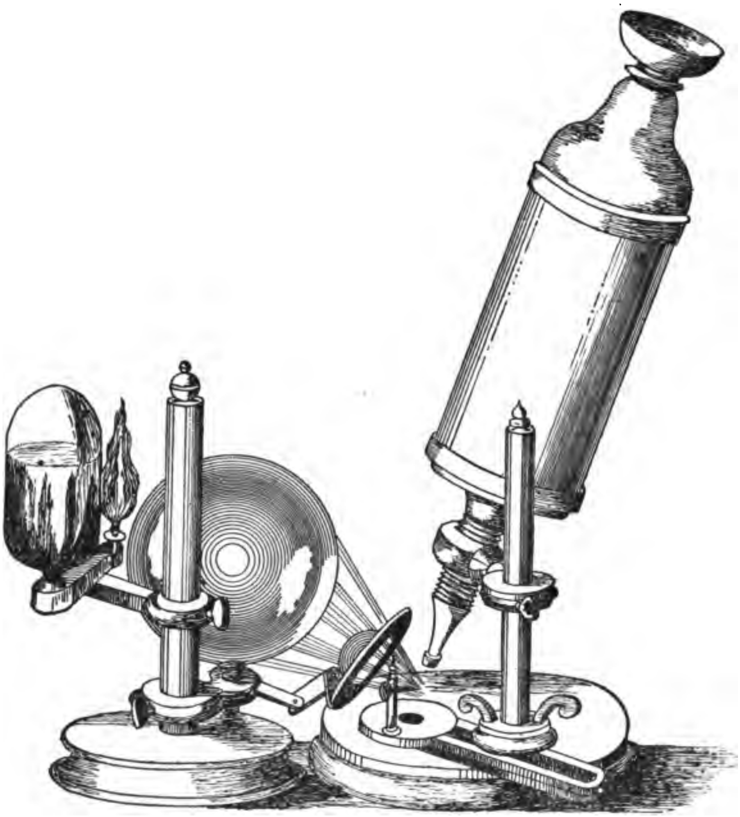

FIGURE 21.

the tube, which served to form a magnified image of an object held at its focus; an eye lens near the top of the tube with which to view the image formed by the objective, and, when he was willing to sacrifice something of the distinctness of vision for an increase of area seen, a third lens between the two mentioned. These three lenses or their

representatives are still characteristic of the compound microscope. The small lens which Hooke called the object-glass is now represented by a system, often very complex, of from two to ten different lenses, acting, however, like a single lens and called the objective. The other two lenses are regarded as a combination, styled the ocular. Of these the one next the eye is called the eye lens, and the other, the function of which is to increase the area or extent of field viewed, is called the field lens. In modern instruments, however, nothing is to be gained by removing the field lens.

Hooke distinctly states that simple microscopes or single lenses of high power yield better vision than his compound instrument, but the difficulty of securing a satisfactory illumination of the object under a lens of very short focus more than counterbalanced their optical superiority. Hooke's method of illumination by means of a lamp and a glass globe filled with water, which acts as a so-called condensing lens, is made clear by the figure, and it is perfectly descriptive of the method used to this day for opaque objects.

Shortly after the appearance of the Micrographia, Leeuwenhoek, a Hollander, commenced the publication of a long series of microscopic observations, which extended from 1673 to 1695. In scientific importance, though not in popular interest, these papers far transcend the earlier writings of Hooke, and were all founded upon discoveries made with simple lenses of his own grinding. Of these Leeuwenhoek left several hundreds, mostly mounted for use and accompanied by an object for observation. His highest powers magnified about 270 times. Twenty-six of these finished microscopes are still in the possession of the Royal Society of London, to which Leeuwenhoek left them in his will. A figure showing the construction of these is given on the following page.

At this point the improvement of the microscope as an optical instrument was arrested for more than a century. In outward form it underwent continuous modification, so that in purely mechanical construction the compound microscope of the beginning of the nineteenth century is not greatly unlike

the best examples of the modern instrument; but leaving out of account the question of greater convenience in use, its absolute power as an instrument of research was less than that possessed by the simple microscopes made and used by Leeuwenhoek.

We are thus confronted by the fact that while the ratio of the improvement in the telescope during the eighteenth century — taking the best telescopes of Huyghens and Campani as

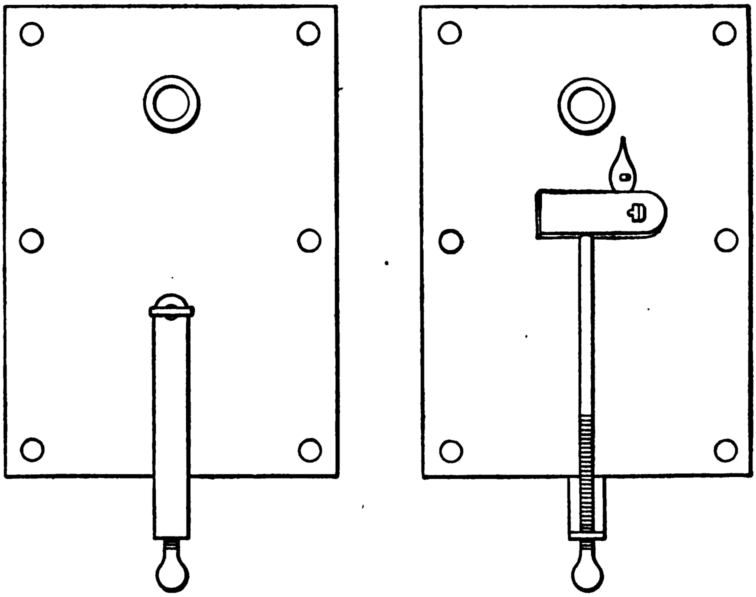

FIGURE 22.

the measure of excellence for the earlier period and those of Herschel and Schroeter for the later — may be reckoned at something between four and six, and the ratio since 1825 at one and a half or two, the corresponding numbers for the improvement of the microscope may be fairly represented by one (equivalent to a stationary condition) for the longer period, and as four for the interval since 1825. This extraordinary inequality in development, perhaps unique in the history of instrumental aids to scientific research, offers a curious

subject for study. Happily the elementary theory of optical instruments given in the preceding pages is quite sufficient to enable us to follow it and to give a rational explanation of many phenomena not generally understood. But it will carry us even further than this; for by its means we shall find a physical limit to the power of a microscope depending upon the very nature of light itself, and thus shall be able to attain a view of what future advances in the art may be expected to bring.

Consider first the simple lens used as a microscope, and strive to find the ultimate limit of its power, either practical, such as may be determined by its minuteness, or theoretical.

The action of a magnifying glass is explained on page 22. Its office is to form a virtual image of an object very near its principal focus at a conveniently greater distance; or, in other words, to render convex wave-surfaces, having their centre at a point in the object, flat wave-surfaces after passing the lens. We shall suppose that this may be done without error, in which case the magnification may be increased indefinitely by increasing the power of the lens. Any increase of power, however, demands increased curvature of the lens surfaces and diminished distance between the lens and the object. For example, a lens of glass of spherical form, which to the unassisted eye would make an object appear one hundred times larger than it would at a distance of ten inches, would require the object to be within one-thirtieth of an inch of the surface of the glass. If the magnification were ten times greater, this distance, called the working distance of the microscope, would be reduced to one three-hundredth of an inch. It is obvious that this fact would put a practical limit to the useful power on account of the difficulties of illumination and adjustment. In special cases the working distance might be considerably increased by using a material of greater refractive power than glass, such as diamond, sapphire, or garnet; and in the early part of the century many experiments were made with these substances, but without the gain hoped for by their advocates.

By the absolute length of the waves of light, however, there is a much more effective limit set to increase of magnification. For it is self-evident that since this element depends upon increased curvature of the refracting surfaces, to increase the power of a lens, we must in the end decrease its diameter; hence the diameter of the wave-surfaces after passing the lens must ultimately be less with greater magnification. Now on page 38 it was shown that vision through a hole much less than one-sixteenth of an inch in diameter becomes notably impaired, because such dimensions are not *very* great compared to the length of a wave of light. From this it follows that the minute details of an image could not be fully recognized if the diameter of the lens were much less than one-sixteenth of an inch. This assertion may be rendered convincing if we consider two points in the object very near each other. The image of each of these points will appear as a disk of a determinate size, and if the apparent separation of the points is not greater than the diameter of these disks they cannot be seen as two points, but as one only. Increasing the power of the lens does not help in the least, for, although it increases the apparent separation of the images, it at the same time, on account of necessarily diminished diameter, increases in exactly the same ratio the diameter of the disk which represents the image of a point on the retina.

The experiments just referred to also prove that when the aperture of the pupil is reduced to one-thirtieth of an inch the indistinctness due to the cause under discussion becomes very obvious; consequently a lens smaller than this in diameter can no longer add to the power of vision, since each point in the image appears as a disk and each line as a stripe, of which the diameter and thickness increase directly with the magnification. But the highest power which a lens of one-thirtieth of an inch in diameter can have is 600,[1] which may

[1] The proof of this statement is given in Appendix A. It is tacitly assumed that there is air somewhere between the object and lens; the modification necessary in case of an optically denser substance surrounding the object will be sufficiently clear from the discussion of immersion objectives.

therefore be stated as the theoretical limit of power for a simple microscope. This conclusion is quite independent of the material of which the lens is made. In every case the practical limit would probably be found considerably below this; at least it is tolerably certain that no discoveries have ever been made with simple microscopes magnifying more than 250 to 300 times.

In the case of a simple microscope the difficulties arising from too close an approximation of the eye and object to the lens can be obviated by the compound microscope; for in this case the eye is removed by a little more than the length of the tube from the object, while a deficiency of power in the objective can be compensated by increase of power in the ocular. Thus there is no necessary relation either between the working distance, or the power of the objective, and the total magnification. It is this feature which led Hooke to prefer the compound instrument to the optically superior single magnifier. However, notwithstanding the ease with which the magnifying power could be increased, it was found that nothing was gained by increasing it beyond 200 diameters. This remained true until later than 1825. The cause of this failure in higher powers was not understood, but it was attributed to defects of refraction at spherical surfaces, which we know under the names of spherical and chromatic aberrations. Even Dolland's brilliant discovery of a method of eliminating the latter error failed to bring any improvement in the microscope for more than seventy years after its publication.

To comprehend the cause of this stationary period of more than a century in the development of the compound microscope, as well as to trace intelligently its later progress to a condition approaching perfection, we must study more critically than appears in Chapter II. the theory of the instrument. To do this we shall gain much by considering the construction from a somewhat artificial standpoint. Instead of regarding the objective as the lens of a camera, as in that chapter, and the ocular as a magnifying glass by means of

which the image formed by the objective is observed, we shall regard the instrument as a simple microscope placed just in front of a small telescope. From this standpoint the objective forms an enlarged image of the object at its principal focus, which is at a great distance beyond the object, and this image is looked at through the telescope. It must not be concluded that this artificial method of consideration is necessary for the establishment of the general laws which we shall derive, but it is convenient because the physical theory of the telescope has already been developed, and by this means we shall escape the necessity of establishing any new principles.

The magnifying power of the compound microscope thus regarded becomes equal to the power of the simple microscope, which constitutes the anterior system multiplied by the power of the telescope. Thus, as far as merely geometrical considerations are involved, it is a matter of indifference whether the desired magnification is secured by the one or the other of the two parts. Let us suppose for the present that the portion which acts as a simple microscope is absolutely perfect, leaving until later the consideration of a departure from this assumption; then the simple microscope forms a perfect image of the object at an infinitely great distance, the apparent size of which is $10p$ times the apparent size of the object when held ten inches from the eye, p being its power. For example, if the power of the lens is 1, that is, if it has a focal length of one inch, the above value is 10; if the focal length is one-half an inch, giving a power of 2, the value becomes 20, and so on. This is a merely numerical extension of what has been stated on page 21.

In the chapter concerning the telescope it was shown that the highest useful power is about thirty times the number of inches in the diameter of the objective. In the case under discussion the available diameter of the telescope is obviously the same as the diameter of the back surface of the lens or lenses acting as a simple microscope. Let this

diameter be represented by $2FN$, F being the focal length of the simple microscope and N a number to be determined later; then, since the total magnification is equal to that of the simple microscope multiplied by that of the telescope, the highest useful power is equal to $10p \times 30 \times 2FN$, or to $600N$, as appears from the fact that p and F are reciprocals of each other.

The largest possible value for N, when there is air anywhere between the object and the front of the microscope, is 1;[1] consequently in the class of instruments under discussion the useful magnification is limited to 600 times. The two obvious and interesting features of this conclusion are, first, that the ultimate useful power attainable with a compound microscope with an air objective is the same as that of a simple microscope; second, that the ultimate useful power of a compound microscope depends upon the ratio of the effective diameter of the rear surface of the objective to its focal length, and not at all upon the power of the objective nor upon the length of the instrument. This last deduction recalls the somewhat similar rule governing the ultimate useful power of a telescope.

The true efficiency of a perfect microscope is thus determinable from the value of N alone, which should therefore be used to characterize an objective. It is called the numerical aperture, and it is generally represented by microscopists by the symbol $N.A.$ Why this awkward term should have been chosen by Professor Abbe, to whom we owe its introduction, will appear later.

Before proceeding with the history of the evolution of the modern microscope, we may perhaps define more closely the term "highest useful power," particularly as it leads us directly to an expression for the maximum defining power of a given type of microscopes. It will be observed that the number 600 in the foregoing expression for power depends

[1] The proof of this important statement is given in Appendix A; it constitutes the foundation of the physical theory of the microscope.

upon our arbitrary assumption of a limit of 30 diameters to the inch aperture in the useful power of a telescope; consequently, it may be argued that the limit of power thus determined for the microscope is also arbitrary. This is true, but not important. What is meant is, that with an absolutely faultless microscope with a dry, or air, objective, all the details of an object visible with any power whatever could be seen with a magnification of 600, and that no one would find any advantage in employing a higher power; most observers, indeed, would prefer 400 or 500. For example, let us suppose that we have at command a perfect dry objective of one-inch focal length and greatest possible value of N, then the available aperture of the telescope, which combined with this objective makes a compound microscope, is two inches. As a result from both theory and practice, however, we find that the closest points or lines which can be seen through a telescope have an angular separation of $4''.56$ divided by the number of inches in its aperture (see page 72), that is, for this case $2''.28$. Now an arc of $2''.28$ is 2.28/206265 of the length of the radius with which it is described; consequently this fraction represents the portion of an inch which must separate the two points or lines that can be just seen as not single by means of an objective of one-inch focal length. The fraction is equal to 1/90000 part of an inch; hence, since the separating power has been shown to depend upon the value of N only, we may make the general deduction that the finest structure which can be resolved in white light by means of an air objective is represented by 90,000 repetitions per inch. If blue light is used, the defining power increases in the inverse ratio of the length of the light waves. For example, if light of the same wavelength as that of the line F in the solar spectrum is employed for illumination, the number rises to 100,000, which may be regarded as the maximum for vision. For a photographic plate, however, the ultimate defining power must be somewhat increased over this limit, possibly twenty per cent.

As to the magnification necessary to exhibit this fineness

of structure, we have the following considerations to guide us: A good eye will just recognize a system of lines separated by intervals of 60 to 70 seconds of arc; but 30 times 2.28″ is 68.4″; hence a magnification of this amount in the telescope for the example chosen, corresponding to a total power of 300, will just enable a keen eye to see the structure, while twice this value will certainly quite reach the limit imposed by the length of light waves even with inferior eyes. This, then, is the meaning of the ultimate useful power being $600N$.

Until the importance of the quantity N was recognized, or at least suspected, rapid improvement in the microscope was impossible. While deficiency in power was attributed to spherical and chromatic aberrations alone, without any effort being made to investigate the real importance of these defects in absolute measure, nothing was to be expected in the way of a rational progress. This was the condition of the affair throughout the eighteenth century. It is true that during that time a method of eliminating one of the supposed barriers to progress was discovered, namely, the principle of achromatic compensation; and this was early applied to the microscope, but without any improvement whatever.

It seems to have been an accident dependent upon the mechanical difficulties of construction, which finally let in light upon the hidden principles involved. It had been noticed as a practical fact that it was by no means a matter of indifference whether the magnification was obtained by a powerful objective or by means of the remainder of the optical system, but that powerful objectives always yielded relatively better results. When achromatic combinations were tried, however, it was found impossible to make very powerful lenses on account of their minuteness, the power of the combination being only the difference of the powers of the positive crown lens and the negative flint lens combined with it. But in 1824 Selligue had the happy thought of combining a number of binary achromatic lenses so as to act as a single lens of increased power and thus retain a diameter

of the system which was as great as that of a single achromatic. His objective, constructed by Vincent and Charles Chevalier,[1] consisted of four like binary lenses which could be used singly or combined as a system of two or more. The focal length of each was about one and a half inches. The following year the Chevaliers improved this construction by inverting each binary lens so that the flat sides are turned toward the object. This notable improvement, apparently so insignificant, was followed by Professor Amici of Modena, in 1827, whose success is emphasized by Lister. Two years later Utschneider and Fraunhofer were supplying objectives like those of Amici, except that the powers of the binary lenses increased from the rear.

In 1830 a very remarkable paper by Lister appeared in the Transactions of the Royal Society, in which he showed that the relative distances of the binary lenses play a very important rôle in the function of the objective. Figure 23, *a*, shows the form of his objective, which is still universally employed for low powers; and shortly afterward, by an extension of the principle, the optician Tully succeeded in producing a triple objective, formed of three binary lenses like the two figured, which possessed a value of N equal to 0.42. Lister's paper also contained the description of an apparatus, the first of its class, for determining the largest angular extent of a wave-surface which an objective can transmit.

This paper by Lister marks an epoch in the history of the microscope. Although its theoretical importance has been greatly over-rated by writers on the microscope, its practical value at the time of publication hardly admits of over-estimation. The immediate improvement was not so great as that made in Selligue's objective, since objectives with a value of N as great as 0.3 were then in use in England; but for the first time a guiding principle was laid down and a method given for determining a constant that bears a

[1] Here, as in all following diagrams of objectives, the object is supposed to be below the lower line of the figure. The shaded areas will always represent sections of flint lenses.

simple relation to the true measure of power which is here indicated by N. From that time progress toward the high degree of excellence recently attained has been continuous. Dr. Goring had already introduced the practice of determining the efficiency of a microscope by means of "test objects," such as the finely lined scales of various insects and, with

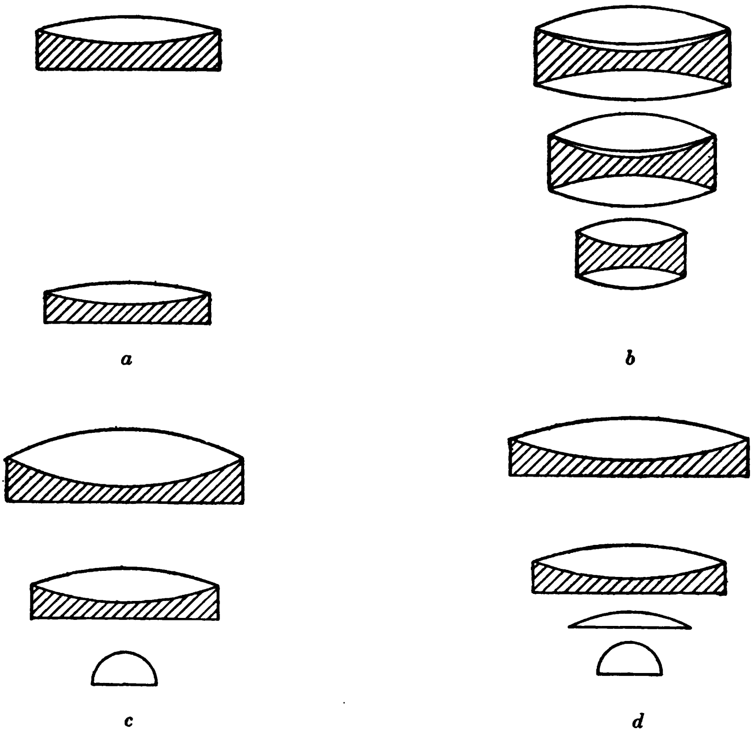

FIGURE 23.

later and better microscopes, the delicate silicious shells of many varieties of diatoms. These tests and Lister's constant, which soon received the name of "angular aperture,"[1]

[1] The true measure of efficiency, which has been indicated by N in the preceding pages, is in fact the sine of one-half of Lister's constant. It was doubtless in order to distinguish the new term, and at the same time to suggest its relation with the old one, that Professor Abbe employed the somewhat infelicitous expression "numerical aperture."

remain to this day the ordinary means by which microscopists and makers characterize their objectives.

Before proceeding with this sketch of the history of the growth of the microscope it is well to note the work done in this country by a young man, Mr. Edward Thomas, more particularly since it has been quite overlooked by writers on the history of this instrument. In the "American Journal of Science" for 1831 Mr. Thomas has two papers describing, with all the detail necessary to guide a practical optician in their construction, two objectives, the more powerful of which is represented diagrammatically in Figure 23, *b*. This objective had a numerical aperture of 0.6 to 0.7, and was doubtless the best that had been constructed up to that time — a supremacy which it retained for a period of ten or fifteen years. Since this writer seems also to have had a deeper insight into the nature of the difficulties to be surmounted in perfecting the microscope than his contemporaries, we may fairly conclude that his death in this same year was a real loss to science.

Aside from this forgotten achievement no advance in achromatic objectives is recorded between 1830 and 1837. At least, Goring and Pritchard in their "Micographia," published in the latter year, regarded the reflecting microscopes constructed on Amici's plan, with mirrors made by Cuthbert, as superior to any achromatic at that time attainable. Probably no man then living was a better critic of the optical excellence of a microscope than Goring, and, as the greatest value for N with his instruments was 0.46, we may safely accept that number as a measure of the highest efficiency then reached. It is worth noting that only at this epoch did the reflecting microscope ever rival the refracting, and even then Cuthbert alone was able to make the essential elliptical mirrors of such great aperture.

In 1837 Andrew Ross introduced the practically important improvement of so mounting the lenses that the distance between the front and middle combinations could be slightly changed at will, thus compensating more or less perfectly the

altered conditions attending different thicknesses of cover-glasses. This adjustment is indispensable in dry objectives of large aperture. During the decade following this, Ross and Amici were leaders in improving achromatic objectives, although at the end of this period Powell and Smith in England, Oberhaeuser in France, and Spencer in America were almost, if not quite, on an equal footing with their predecessors. Oberhaeuser and his successor Hartnack deserve special mention, because their efforts to produce efficient microscopes at a relatively small cost were so far successful that doubtless more serious scientific work has been accomplished with instruments from their factory than with those from any other source. To this period belongs a type of construction for objective-fronts, introduced probably by Ross, which remained for some time the prevailing construction for high powers. It is shown in Figure 24.

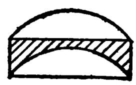

FIGURE 24.

Between 1850 and 1860 was introduced the practice of making the anterior lens of high-power objectives of a single piece of crown glass nearly hemispherical in shape. Although the value of this innovation stands unapproached in the history of the microscope, excepting possibly those of Selligue and of Lister, it is impossible to name definitely its inventor. It is generally attributed to Amici; but Wenham, whose contributions to practical microscopy have been important in other fields, states that he introduced it in 1850 and exhibited it in successful employment. The form in which this improvement was rapidly adopted by leading makers of that period is shown in Figure 23, *c*, while an extension in the same direction, indispensable for the highest attainable powers, and probably invented by the skilful American optician Tolles, is exhibited in Figure 23, *d*; it is known as the double front. But whosoever the discoverer may have been, the discovery itself was not only a necessary step in the progress toward perfection, but, once taken, it led inevitably, by a rapid approach, to modern excellence.

The theory of the hemispherical front is not difficult to

comprehend. The office of every lens in an optical instrument is to alter the curvature of the wave-surfaces transmitted by it; if the transmitted wave-surface is of constant curvature, that is, either flat or strictly spherical in form, the lens is free from spherical aberration; and leaving out of consideration the chromatic aberration which can be corrected by other means, the resulting image is geometrically perfect. Now in general a spherical refracting surface does not accomplish this, although, with the present mechanical means at command, spherical surfaces alone admit of precise shaping. The only practical method of attaining the required end is to combine convex and concave surfaces so that their errors, opposite in kind, shall, after the method of Lister, compensate each other as far as possible. The resources of this method, however, are exhausted at about the limit of $N = 0.6$, that is, further progress requires either a new principle or a method of making accurate lenses of other than spherical forms. The latter means appears no less hopeless now than in the time of Descartes. Fortunately a new principle was discovered in the utilization of one of the two special cases in which refraction at a spherical surface is geometrically perfect. This case may be explained by reference to Figure 25, which represents a piece of glass bounded at the left by a polished spherical surface whose centre is at C. A line Cp_1 through the centre is chosen as the axis of the lens. If any point of this line more remote from the surface than C is a source of light waves, they will suffer such a modification in passing from the glass into the air as to reduce their curvature, although in general they will also lose their spherical form. It is obvious that if the source corresponds with the centre C, neither the curvature nor the shape will be changed; but it is also true that there is another point p_1, from which spherical waves will be refracted without loss of their spherical form, and at the same time undergo great diminution in curvature. The position of this remarkable point is defined by the relation Cp_1, equal to the radius of the surface divided by the index of refraction of the glass. A very simple computation

will show that the image of this point p_1 is at p_2 so placed that Cp_2 equals the radius of the refracting surface *multiplied* by the refractive index; consequently the ratio of Cp_2 to Cp_1, which is also that of the magnification of the image, as is obvious from the figure, is equal to the square of the index of refraction, or to n^2, if we employ the ordinary symbol for that constant.

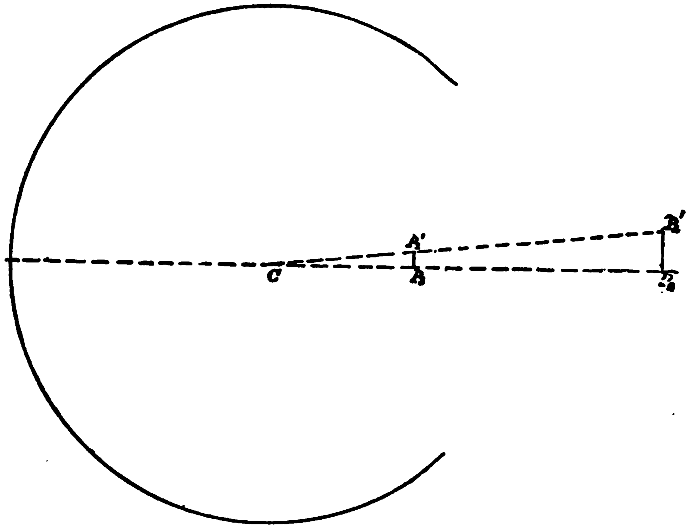

FIGURE 25.

Here, then, is a means by which an object inside a sphere of glass and a short distance from the centre can be replaced by a virtual image n^2 times as great in size, and absolutely without faults except such as depend upon the small variations of n for different wavelengths. We may first turn our attention to the limit of the advantage to be thus acquired, and afterward consider the practical methods of securing the proper physical conditions.

In Figure 26 let p_2 represent such a virtual image of p_1, magnified n^2 times, and *aob* an indefinitely thin lens system everywhere equidistant from p_2, which will render wave-sur-

faces from p_2 exactly flat. It is evident that the whole system is just that which, combined with a telescope, will constitute a compound miscroscope in the sense of the previous analysis. Moreover, it must be admitted that this highly artificial construction is at least as efficient as any other, because by assumption each lens performs its office perfectly.[1] But the elements of the system are so simple that it is quite easy to calculate the ultimate magnifying power. Thus, in the expression for the highest useful magnification on page 93, the value of the power is here equal to the product of the two factors, namely, n^2 due to the first refracting surface and $1/op_2$ due to the lens system ab; again, the aperture of the telescope is equal to the diameter of ab, or to $2op_2 \sin ap_2o$. Substituting these values in the expression, there results:—

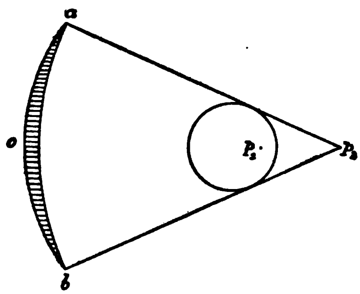

FIGURE 26.

$$600n \sin ap_2o.$$

The highest possible value of sin ap_2o, as appears almost directly from Figure 25, is $1/n$; hence the value of the ultimate useful power of any microscope constructed on this system is

$$600nN,$$

N, as before, being a number never greater than unity.

From this analysis it appears that the range of microscopic vision is only limited by the refractive index of the material of which the front lens is made. Unfortunately very high refractive power seems to be inseparable from opacity, and we know nothing now which can approach diamond as a possible material, though it would obviously offer great mechanical difficulties in use: its index of refraction is about 2.5. Many other substances of very high

[1] See Appendix A for a further discussion of this point, and also of the maximum value of ab.

refractive powers, such as sulphur and phosphorus, are rendered impossible in use, either because of double refracting property or of their mechanical condition; while the densest glasses, with indices approaching 2, have such large dispersion that their employment would introduce complications most difficult and perhaps impossible to deal with. Ordinary crown glass, with its moderate index of 1.5, appears, therefore, to be the best substance for practical use. An objective with a front lens of such a material and the object within its substance should enable us to see lines as separate objects when as close as 135,000 to 150,000 to the inch. Since our first microscope-makers, under the leadership of Professor Abbe and Dr. Zeiss, have attained within five or ten per cent of this limit, it seems not extravagant to speak of the microscope as having reached practical perfection.

The methods by which in practice the conditions involved in the preceding theory are secured may be indicated in Figure 27. Here the hemispherical fronts of three objectives are

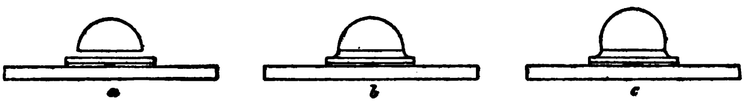

FIGURE 27.

represented with the objects to be viewed below them under a cover of thin glass. In each case the virtual image of the object, formed by all the successive refractions suffered by the wave-surfaces up to that by the spherical surface itself, must be very near the point corresponding to p_1, in Figure 25; but the continuity of the glass is interrupted by a thin plate of fluid which admits of the necessary adjustment of focus, the fluid being air in *a*, water in *b*, and in *c* a liquid as similar as possible in its optical properties to those of crown glass. The introduction of these layers, save in the last case, causes errors in the refractions which become greater with greater thickness and wider difference between their optical properties and those of glass. Thus, in the first case, where air

is the dividing medium, the layer must be very thin if a large aperture is to be employed, or errors will be introduced which do not admit of correction in the remainder of the system; in short, the "working distance" of an objective of high efficiency of this type is necessarily very small. It is ordinarily called a dry objective. An objective constructed with a front such as is illustrated by *b*, the water immersion objective, possesses much greater flexibility in respect to working distance, because its conditions are much nearer those demanded by theory. Finally, the homogeneous immersion objective, often called the oil immersion because the fluid employed is the oil of red cedar, may have a working distance up to nearly two-thirds the radius of the hemispherical front and still perform satisfactorily.

It is easy to show that the expression given as a measure of the ultimate power of the homogeneous immersion objective, namely, $600nN$, would apply to all forms if n represents, not the refractive index of the substance of the front lens, but that of the least refractive medium between the object and the spherical surface of the front lens. Thus, for a dry objective $n = 1$, for a water immersion $n = 1.33$, and for a homogeneous immersion $n = 1.5$. As for N, it is as easy to obtain a high value in one type as in another, and the most skilful makers succeed in constructing excellent objectives with a value as high as .93 to .95. From these data and the discussion which appears on page 93 we may conclude that the limit of resolution of dry objectives is about 90,000 lines to the inch, of water immersion about 120,000, and of homogeneous immersion objectives about 135,000. It is an interesting fact that in order to secure this limit in the type of objectives last named, it is necessary to employ more than a full hemisphere in the front lens, as is evident from an inspection of Figure 25. This obviously heightens enormously the difficulty of shaping and mounting such lenses; hence all objectives of the greatest possible aperture must be costly.

No review of the microscope at this period, however elementary its aims, can leave unnoticed Professor Abbe's very

remarkable invention which he has named the apochromatic objective. To give a rational description of this will require, it is true, a step further in the theory of optical instruments than has been necessary up to this point, although it is a step which follows very naturally what precedes; in short, it will be necessary to consider some of the chromatic defects which are inseparable from the old construction. For this end, suppose the objective, which, in accordance with our previous convention, serves only to make a perfect virtual image of the object at an infinite distance, to be divided into

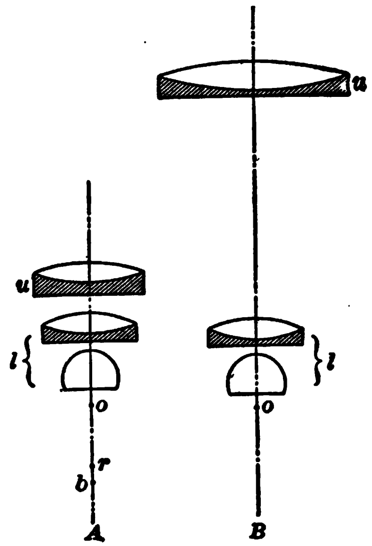

FIGURE 28.

two portions, the lower of which (*l*, Figure 28, *A*) forms a magnified virtual image of the object at a finite distance, and the upper part *u*, a corrected virtual image of this at an infinite distance. But since all transparent substances refract short waves more than long, the image of *o* formed by *l* will be a complex of colored images from *r* to *b*, the blue image being the most remote.[1] The function of the system *u* is to con-

[1] Of course the system *l* is undercorrected for color; if not, let *l* apply to the lowest lens alone, and *u* be a system comprising the two binary lenses, and the reasoning will be in no wise changed.

vert all the complex system of wave-surfaces having their centres distributed from *r* to *b* into plane wave-surfaces; in other words, the system *u* must have a higher power for red than for blue waves, and at the same time be sensibly free from spherical aberration for all colors. There is no difficulty in meeting these conditions for wave-surfaces of moderate angular extent by a combination of negative lenses of flint glass with positive lenses of crown glass, but when the wave-surfaces are large, as they must be in powerful objectives, it is found that the power of the upper system is always relatively too small at the margin for the short waves of light. This want of flexibility in the means of correction arises from the fact that in all known glasses a great increase in dispersive power is invariably accompanied by increase in refractive power. It is true that there are many liquids combining high dispersive power with relatively small refractive indices, and by a combination of such a liquid with crown glass and flint glass in the system *u*, Professor Abbe in conjunction with Dr. Zeiss succeeded in constructing two objectives of unprecedented excellence; but such fluid lenses are practically quite useless on account of the mechanical difficulties of keeping them in working order.

Fortunately for science the elimination of this most serious defect in high-power objectives did not depend upon the discovery, not very promising it must be confessed, of a variety of glass which should possess the optical properties of some exceptional fluids. Professor Abbe devised a brilliant expedient, which seems to have occurred to no one before him, by means of which this defect is eliminated in an entirely different manner. Consider the changes in the problem if the lens system *u*, instead of standing as close as possible to *l*, is removed to a considerable distance, as in Figure 28, *B*. Here we recognize that the change of curvature to be produced by *u* is lessened, and hence that the power of the objective as a whole is somewhat decreased; but what is vastly more significant is that the *difference* of curvature between that of the red and that of the blue is decreased in

a much greater ratio. Thus, if the curvature of the red light waves is half as great in the second case, this difference will be only one-fourth as great; if the curvature is reduced three times by making the distance between the lens systems *l* and *u* twice as great as that separating *l* and *r*, this difference will be reduced nine times, and so on. The modification consequently admits of varying the relation between the two distinct functions of the system *u* within very wide limits, and thus yields another arbitrarily variable element for the attainment of a closer correction. A rational use of this principle with a skilful choice of materials, so as to reduce to a minimum the far less serious defect of what is called secondary chromatic aberration, yields the highly refined apochromatic objective.

The construction of the apochromatic entails a defect of a singular character in the complete instrument, that is, when used as a part of an ordinary microscope, which is of interest on account of its general nature. In Chapter II. it has been shown that the magnification of a virtual image depends upon the change in curvature of the wave-surfaces originating in the object; hence the lens systems *l* of Figure 28 magnify more for short wavelengths of light than for long, and consequently the whole objective will magnify a blue object more than a red one unless the systems *u* exactly reverse this relation, that is, unless they change the curvature of the wave-surfaces of short wavelength less than those of longer waves, and less by just the proper amount. Now, whatever may be the case with the system *u* in Figure 28, *A*, it is clear that the corresponding lens system in Figure 28, *B*, cannot correct the defect in question in the anterior portion, because the separation has been made for the express purpose of reducing the relative difference in the curvatures of the wave-surfaces. Thus it appears that in this objective, and in general in any lens system in which correction for color is obtained by lenses separated by a considerable distance from uncorrected lenses, the images of an object, although all in the same plane, and therefore achromatic in the ordinary sense of the word, differ

materially in magnitude with differing color. This defect produces a characteristic imperfection in the seeing, with all high-power microscopes, which appears as a form of color error, increasing very rapidly with increasing distance from the centre of the field, a phenomenon perfectly familiar to all who have observed critically with such microscopes, although it is not generally understood. Professor Abbe corrects this defect by adding to the ocular a compound lens of such construction that it makes its power greater for red light than for blue in the same ratio as the excess of magnification of the objective lies in the opposite direction. These two elements combined, namely, the apochromatic objective and the compensating ocular, form the modern perfected microscope.

In the theory of the compound microscope there is still remaining one point of more than merely technical interest, and that is the question of distribution of power between the objective and the ocular. The statement has been previously made that the experience of early constructors showed that this question, though not suggested by theory, was a most important one. They found that in general much better results were attained by the employment of powerful objectives and weak oculars. The reason for this is not now difficult to explain. Ignoring the small but inevitable errors of construction, we have just learned that there are other unavoidable sources of imperfection in the objective, arising in part from the fact that errors of refraction cannot be corrected where they occur, but only in more or less remote portions of the system, which must bring with them imperfections in the images to be enlarged by the ocular. How small these faults can be made is a question of experience rather than theory, because of their enormous complexity in kind and origin. Such experience has shown that with the ordinary type of powerful objectives the images are so far from perfect that they will not bear a magnification of more than five or six times without passing the limit of definition imposed by the inherent faults of construction; in the more perfect apochromatic objectives this magnification may be

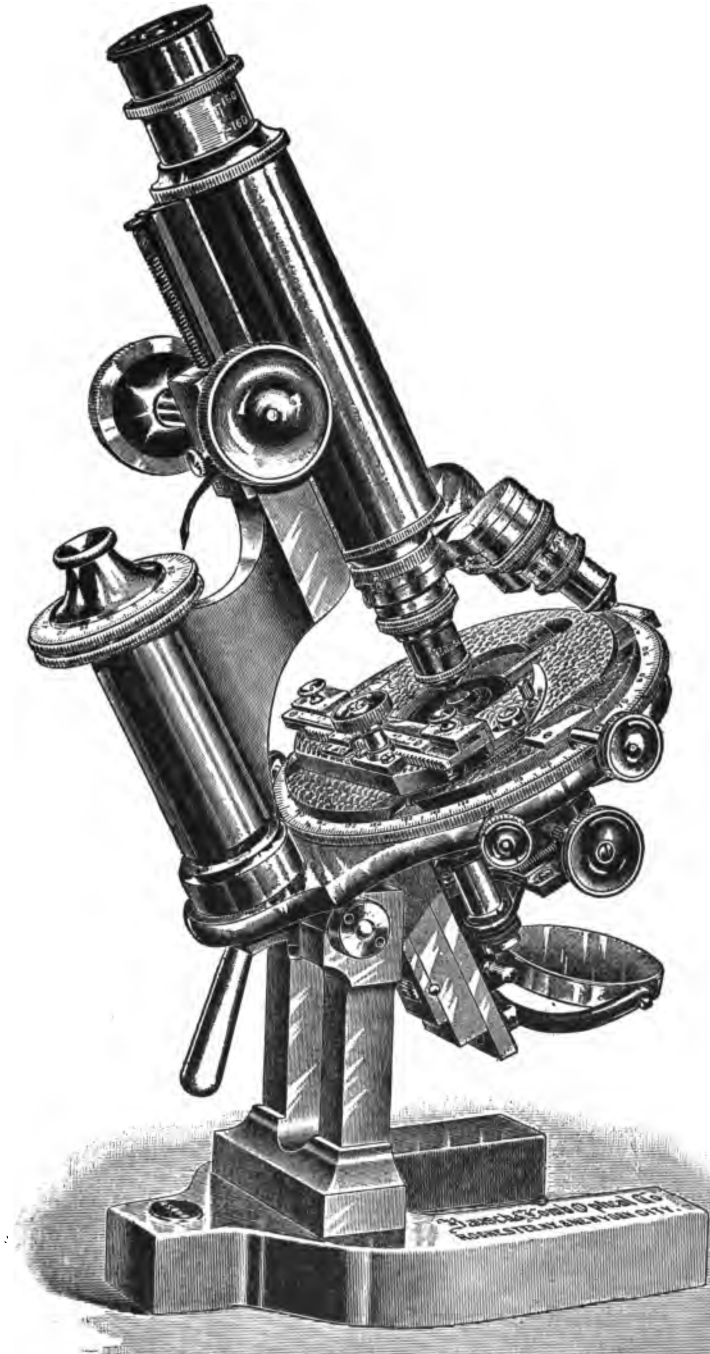

Figure 29.

raised as high as from twelve to fifteen. As we have found that the highest useful power of a dry objective is about 600, we conclude that a good objective of the old type of a power of 100, that is, of 1/10 inch focal length, or of the apochromatic construction of 1/5 or 1/6 of an inch, will accomplish all that any other, however powerful, can. In a similar way we are led to conclude that the older type of homogeneous immersion objective of 1/15 inch focal length, or an apochromatic of 1/10 inch, are quite sufficient to exhaust the range of microscopic vision, at least as far as our present knowledge of the nature of light and of vision enables us to declare.

On page 86 appears a figure of the earliest compound microscope of which we have much knowledge. The accompanying picture (Figure 29) of an instrument which may be regarded as one of the best existing models serves as an interesting contrast. It is provided with three objectives of different powers, any one of which may be brought into use at once by a rotation of their common carrier. The apparatus below the stage is for the purpose of securing at will illumination from any desired direction and angular extent. It is known as Abbe's illuminator.

CHAPTER VII

OPTICAL PHENOMENA OF THE ATMOSPHERE

THERE are many phenomena, some beautiful, some merely curious, which depend upon the modifications that light undergoes in its passage through the atmosphere. In certain cases these phenomena are necessary consequences of the optical properties of the air alone, or, at least, dependent upon invariable constituents of the air, as, for example, the blue color of the sky, mirages and looming, and scintillation of the stars; in other cases we find the causes in bodies temporarily suspended in the atmosphere, as in the rainbow, corona, and halo.

If the atmosphere were absolutely transparent, that is, if light waves could be regularly transmitted through it without loss, we should find the sky quite black except where a bright spot marked the presence of a star or planet. In short, the sky of day would differ in appearance from that of night only by the presence of the sun. On the other hand, if the air should transmit but a portion of the wave-energy, converting the remainder into heat or some other form of energy inappreciable by the eye, we should have no closer a resemblance to the sky of experience; it would still remain black except at points in the direction of stars a part of whose light would reach the eye. As a matter of fact, Professor Langley has demonstrated that what we call a clear atmosphere is only slightly opaque to light waves. The atmosphere is then neither a perfectly transparent body nor an imperfectly transparent one. What it is we may possibly picture to ourselves most quickly and accurately for our purpose in the following way: Imagine a perfectly trans-

parent atmosphere — therefore a black sky — with a high sun; now imagine distributed throughout this atmosphere small drops of water, say one-thousandth of an inch in diameter. If these are infrequent, perhaps one to each thousand cubic feet of space, the sky would send light to the eye from all directions, since each drop would scatter the light which fell upon it; notwithstanding this action, the brightness of the direct sunlight might not be notably diminished. Such a sky would be called a hazy sky. If the number of drops should be continuously increased, the density of the haze would grow while the brightness of the direct sunlight would progressively diminish until the sky became entirely overcast and the sun invisible. With the exception of the initial black sky, just such a change has been repeatedly observed by every one. Now reverse the process, that is, suppose a portion of the drops to be removed, leaving the distribution uniform, until the sky is no brighter than a clear sky as we ordinarily observe it; then suppose the drops to be reduced in size but at the same time increased in number at such a rate that the total quantity of light from the sky remains unchanged. This reduction in size has introduced an entirely new element into the consideration, for when this is carried so far that the particles of water have a diameter which is small compared to a wavelength of light, they are no longer capable of reflecting these waves, precisely as a floating body on the ocean would be incapable of reflecting waves whose lengths are great compared to its own dimensions, although it would be a perfect barrier to the passage of short waves. Before reaching this condition of extreme tenuity, however, we should have passed through a range of dimensions which might be called small when measured by the lengths of red light waves, but not small when measured by lengths of blue or violet waves. Such particles would reflect violet light and blue light more copiously than orange and red lights. This is the explanation of the blue color of the sky.

It is immaterial what the particles are made of, provided

that they are sufficiently small and not coagulated; hence, when we receive light diffused from a cloud of smoke which is not too dense, especially if the background is black so that no other than this diffused light reaches the eye, we see a similar pale blue. The blue of the opal and of opalescent bodies has its origin from a similar cause, as has also the color of blue eyes and of the deep sea. It is well known to painters that generally a mixture of a black with a white paint gives a strongly bluish gray, although there may be no suggestion of this color in either of the components. This, too, is explained by this species of selective reflection depending upon minuteness of reflecting particles.

From the preceding considerations it follows that sunlight which has come to us through the atmosphere has lost more in short than in long waves; consequently the hue of such light is somewhat yellow. If the light has passed a very long distance through the air, as when the sun is near the horizon, we may have, together with a very great diminution in the strength of all waves, a practically complete stoppage of the short waves. This would leave yellow, orange, or red, depending on the completeness of the action, and, also (as we shall see when we come to study the phenomena of color sensations), to a considerable extent upon the absolute intensity of the light. The colors are strongest after the sun is below the horizon and sends light to us only by the medium of reflecting clouds, for then the path of the light through the air is much longer than when the source is above. The bluish greens and blue-greens which are not infrequently seen in a sunset sky are often a physiological effect of contrast.

Quite an analogous phenomenon to the last is presented by a light smoke which appears blue against a dark background, but yellow when the background is such as to send much more light to the eye than the smoke itself.

One of the most noteworthy effects of this peculiar opacity or opalescence of the atmosphere is the change it produces in the aspect of distant objects. Thus two surfaces, the one

light and the other dark, lose something of their contrast as they recede from the eye, the first becoming darker by the absorption of its light by the intervening air and the second brighter from the superadded light diffused by the air. If the air is free from relatively large particles of water and from coarse dust particles, this added light is blue, and it often requires only a moderate distance to give a strongly blue hue to shadows on a sunlit rock, while distant hills are always of a strong violet or blue tint in a clear day. This effect goes under the general term of aerial perspective, and it affords the readiest means of estimating the distances of remote objects on land. In an exceptionally clear and dry atmosphere the effect is greatly diminished, whereas a moist climate enhances the impressiveness of mountain scenery. A light fog or haze often lends a charm to landscapes which otherwise would be quite uninteresting.

There is another property of the atmosphere, which, though sufficiently obvious to the astronomer, is ordinarily overlooked, at least in its common manifestations, by one whose attention has not been especially directed to it. The velocity of light waves in air of the prevailing density at the surface of the earth is about three parts in ten thousand less than in a vacuum. From this and the decreasing density at higher altitudes it follows that when light enters obliquely into the atmosphere its course is changed by refraction to an amount, when the obliquity is greatest, nearly equal to the diameter of the sun. The sun, therefore, appears to be just above the horizon when, were there no atmosphere, it would appear to be just below it. The effect is to lengthen the day at the equator by about four minutes, though at higher latitudes this lengthening would be greater, even rising to many hours in extreme latitudes. This atmospheric refraction is attended by the secondary phenomenon of dispersion, as are other cases of refraction, but it requires a powerful telescope to detect this fact, and it is consequently of minor interest.

Far more striking than the regular atmospheric refractions

are the curious phenomena known as looming and mirage, which find their cause in temporary inequalities of atmospheric density. The complexity of these phenomena is enormous, nor, except in most general terms, have they been adequately explained; but after establishing an optical principle of great importance, we shall find it quite easy to deduce a consistent explanation from an ingenious experiment to illustrate their origin, invented by Wollaston.

The principle named may be stated as follows: At any point in a wave-surface which has its origin in a distant body, imagine a small portion cut out, say in the form of a circle; then, if the light is modified in any way so that this circle is changed to any other shape or size, the object seen from the point named will appear changed in a manner exactly recip-

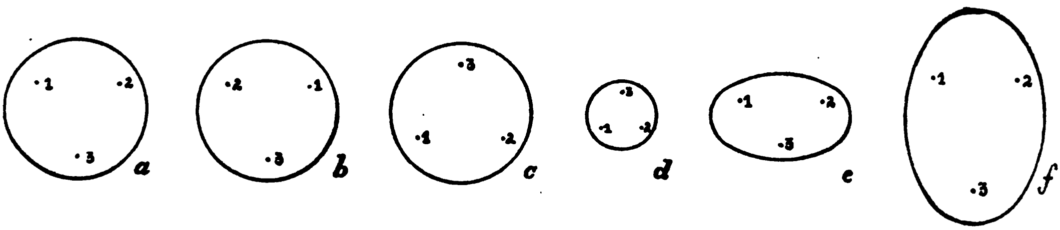

FIGURE 30.

rocal to this.[1] For example, if the circle *a*, in Figure 30, represents the portion cut out of the unmodified wave-surface, in which certain points are numbered for convenience in tracing the changes, and the forms which follow represent respective modifications produced by various optical means, then the first one (*b*) shows that the object is changed from right to left, that is, that it is perverted as it would be by reflection from a vertical mirror; the next (*c*) implies a perversion in a vertical plane; *d*, in turn, shows that the object appears inverted and magnified in the ratio of the diameter of the large circle to the small one; in short, it is this principle which was used on page 70 for determining the magnifying power of a telescope; finally, *e* and *f* indicate, respectively, a

[1] The proof of this important principle is given in Appendix A, section vi.

magnification in the vertical plane, and an equal diminution without change in the horizontal direction.

We now turn to Wollaston's experiment. This consists of a glass tank with parallel sides (he used a square bottle, although it is difficult to find one sufficiently regular to show all the phenomena to be described), into which he poured first a layer of transparent syrup; upon this a layer of water was placed so as not to disturb the syrup; finally, a layer of alcohol was placed with like precaution upon the water. These three liquids are perfectly miscible, but are placed in order of their density, so that they mix only by the slow process of diffusion. Their optical densities, however, follow a quite different order, the middle layer being less refractive, that is, producing less retardation of the wave-surfaces, than either the syrup or the alcohol. If now a distant object in the same horizontal plane as the tank is looked at through the liquid from a point quite close to the tank, the following phenomena may be noted: Starting from a level where the syrup may be regarded as unmixed with water, one sees, first, an erect image of the object in its true position and size; as the eye is raised the object is lifted above its true position and increased in vertical dimension; from a still higher point the object is lifted still more, but recovers its original size; beyond this point follows a less elevation with a diminution in vertical dimension, and, when the eye is at a level where only pure water is in the line of vision, the object is again seen in its true position and magnitude. As the eye is gradually raised to the level of the pure alcohol, the whole series of phenomena is reversed, not only as regards position, but also in respect to elongation or compression in the vertical direction. If the series of observations is repeated, with the eye remote from the tank, perhaps eight or ten feet, much more striking effects will be noted. In the first place the change in level of the eye necessary to bring all the phases into view will be found to be much larger. Secondly, there suddenly appears, at a considerable height above the unmodified image of the object, a second image, which quickly betrays

itself as a double image, the lower of the two being inverted. As the eye rises, the under one of these two images leaves the upper and approaches the lowest of the three, with which it finally unites and ultimately vanishes. For some distance above this point, depending upon the thickness of the stratum of water which may be regarded as unmixed with either syrup or alcohol, only one image will be seen which is unchanged in size only when undeviated; in every case diminution in the vertical plane accompanies deviation. A further raising of the eye carries one through a reversed repetition of the phenomena. The ratio of vertical height of the inverted images to that of the erect ones varies with the distance of the eye from the tank, and is not especially significant for our purposes; but when this distance is the least that admits of a good inverted image, it will be found to be magnified in the vertical direction. What is of more interest to us in its bearing upon the explanation of mirage and other extraordinary refractions is the fact that only from below the lower transition stratum and from above the upper transition stratum can multiple images be seen; in other words, recalling the optical character of the fluids used, only from that side of the transition strata upon which the more refrangible fluid is found.[1] The explanation of the first series of phenomena will appear very simple from a consideration of Figure

[1] The experiment herewith described is as simple to carry out as it is interesting. A tank having a distance of four inches between its plate-glass ends will be found of a convenient size. Into this may be poured a dilute solution of sugar; upon this float a thin piece of wood or cork, and pour slowly a proper depth of water upon the float. This procedure will secure a separation of the two fluids. If desired, the alcohol may be introduced afterward in a similar manner. In order to attain the most satisfactory results the vessel should remain undisturbed for a number of hours, to secure a sufficiently gradual change in density in passing from one region to another. In addition we may note that too great a difference of refractive power is likely to be found if the syrup is not rather dilute, but if two solutions only are used, this is very readily adjusted by thoroughly mixing and replacing a part of the mixture by pure water. Also a gain in distinctness will be found by observing through a horizontal slit of, say, a fiftieth of an inch in width, as this reduces the effect of astigmatism due to the cylindrical form of the wave-surfaces.

31. Here the line AB represents the position of the transition layer which is defined as the region where the ***rate*** of change of optical density is a maximum, and the lines CD and $C''D'$ the position of layers below and above which, respectively, the density may be regarded as constant. A flat wave-surface entering from the right, since it travels faster in the region above the line AB, and since there is no discon-

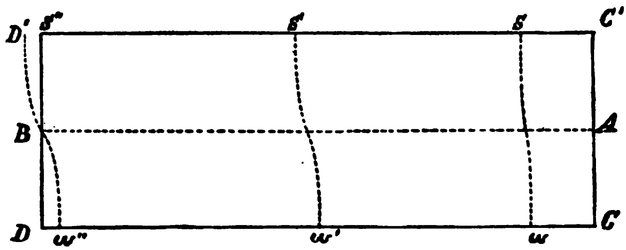

FIGURE 31.

tinuity in the fluid, will take successively the forms represented by ws, $w's'$, $w''s''$. When it emerges from the tank, it will be, in general, slightly modified both in direction and in curvature, but as this change will always be very small and in no case will lose its line of inflection, which is represented by the straight part of $w''s''$, we may ignore this modification in our explanation. If the eye is placed at a point near D, where the wave-surface is flat and perpendicular, the point-source of the wave, and therefore the surface of the object in its immediate neighborhood, will appear in its true place, unchanged in magnitude, just as if seen through any plane and parallel plate. But as the eye is raised toward B, the object will seem to rise above its true position, since the source of a wave-surface will always appear in a direction at right angles to this surface and at the same time stretched out in the vertical direction, as is manifest from the optical principle given above. A little higher the vertical magnification will become less, since the curvature of the wave-surface grows less, but the elevation will increase until

the eye reaches the level *B*, where the elevation is a maximum, yet the dimensions are unchanged because the wave-surface is there again flat. As the eye moves from this point upward, the elevation of the object seems to decrease, but it is subject to a deformation of a contrary character, that is, it will appear to be reduced in its vertical dimensions, this reduction attaining a maximum where the wave-surface has its greatest curvature, and finally vanishing with the elevation at the point *D′*. If higher up there were another transition layer of an opposite kind, such as would obtain if a stratum of alcohol were poured upon the water, all the phenomena described would ensue in a reverse order. We need not, however, stop longer over this particular series of observations, since the very few atmospheric phenomena here imitated can be better discussed when we turn to that series presented to the eye remote from the tank.

We shall find it advantageous for our purposes to consider separately the two cases of the transition layer lying above the less dense medium (alcohol-water) and above the denser medium (water-syrup). Taking these in the order named, we shall meet first with those phenomena of extraordinary atmospheric refraction which are most familiar.

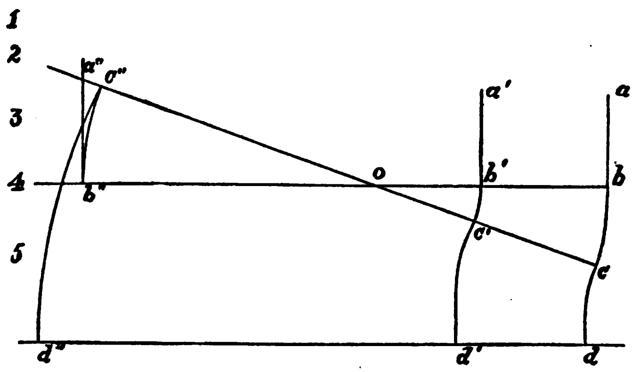

FIGURE 32.

In Figure 32, let *abcd* represent the wave-surface after modification by a refractive medium which possesses a suffi-

ciently rapid increase of optical density; this wave-surface will move on parallel to itself until it may be represented by the line $a'b'c'd'$, in which the straight portion ab remains unchanged, but the concave part bc is greatly contracted, while the remaining convex portion is considerably elongated. Still later the surface assumes the form indicated by the line $a''b''$, or the straight portion, still unmodified; $b''c''$, formerly the concave portion, now rendered convex and inverted, and $c''d''$, which remains convex as before.

This diagram is all that is necessary to explain these phenomena of vision, although it is well to recognize that, in order to represent quantitatively the alcohol-water experiment, the horizontal scale of our figure should be increased a hundred-fold or more; and to extend it to the cases presented us by occasional atmospheric conditions a thousand-fold might be none too little. To an eye in the region *1* of the diagram, only a single image of the distant source can be seen, namely, that due to the flat portion of the wave, and consequently unaltered both in direction and magnitude. At *2*, however, a sudden change takes place, for here three portions of the original wave-surface enter the eye, the straight portion giving an undeviated and undistorted image; a depressed erect image belonging to the convex portion of the wave, also undistorted, since the wave-surface is here practically flat; and, finally, an inverted image superimposed upon this, likewise undistorted. As the eye sinks into the region *3*, the second and third of the images described rise toward the first, but this displacement is more rapid in the case of the inverted image, which at the point indicated by *3* is midway between the others. It is evident from the figure that the erect, depressed image is always smaller in the vertical direction than the undeviated image, but that the distortion of the inverted image depends upon the distance of the eye from the region marked *o*, so that if this distance is too small the vertical magnification will become indefinitely large and the corresponding image will not be seen. This is the reason why in the experiment it is necessary to remove the

eye a considerable distance from the tank, in order to see an inverted image. Finally, when the eye is carried into the region marked *5*, only a single image will be seen slightly depressed and shortened vertically. It is important to observe that the multiple images, whether double or triple, are seen only when the eye is between the points marked *2* and *4*, both of which are above the level of the transition stratum; in short, we may state the rule, which will prove useful later, that only when the eye is on that side of the transition stratum upon which the more refrangible medium lies is it possible to see multiple images. There remains but a single remark to be added before we are in the position to explain the most familiar of all cases of extraordinary atmospheric refractions, namely, that if the eye remains fixed in position and the distant object is moved downward, it will undergo successively all the changes here described. It is easy to see that this follows, from the fact that such a change simply changes the inclination of the wave-surfaces by a small amount (recall in this connection that the horizontal scale of the figure should be increased hundreds of times), which would thus leave the form of the surfaces practically unchanged. Therefore, if one observes a surface indefinitely removed from the tank, all the foregoing phenomena may be seen on this surface simultaneously; and, since all variation of the refractive action is confined to the vertical direction, the resulting modifications in the appearance of the object will be arranged in horizontal bands. For this reason a coarsely printed card at a considerable distance is well adapted for the experiments with the tank.

Although atmospheric air never has mixed with it any substance which, like the sugar in the tank experiment, modifies its refractive power materially, still, since its refractive power decreases with its temperature, it is quite conceivable that in some conditions the temperature might be so distributed that a similar layer arrangement of refractive power might ensue so as to yield many of the results discussed in the pages immediately preceding; and indeed this

is the fact. For example, when cold air lies above the relatively warm water of a lake or sea (a state arising from the appearance of a cold wind, or, more frequently, from nocturnal cooling of the air by radiation), the condition in question is very frequently found. Indeed, in late summer or early autumn, after the waters of our lakes and bays have been accumulating heat for a long time, it is a rare exception when such phenomena cannot be seen in greater or less perfection during the cooler portions of the day. Under such circumstances the air in the immediate neighborhood of the water is warmer, less dense, and less refractive than the air at higher levels; consequently an object at a distance greater than our sensible horizon and sufficiently near the horizon will send light to us which on a part of its path has been subject to the optical modification illustrated in Figure 32. It is true that this light does not enter the region where it suffers this extraordinary modification, through a vertical plate; but this fact, although it complicates the problem, does not alter its essential nature. We shall have, therefore, a more or less complete representation of the phenomena discussed in connection with that figure. There is, however, one difference of material moment in the atmospheric analogue, which must be here noted. The warmer air below is not only less refractive, but it is also specifically lighter; hence it is not in equilibrium, and constancy of the optical images can be hardly expected. On the contrary, the intense unsteadiness of such images, which often results in most grotesque changes in form, is one of their striking peculiarities. It also follows from this instability of equilibrium that not only must there be a constant supply of heat from below to sustain it, but a true transition stratum is never present, since the temperature must be highest just at the boundary surface. From the second consequence it obviously follows that the convex portion of the wave-surface is wholly missing, and hence the erect image below is never seen. By reference to Figure 32 it is now easy to describe the phenomena which would follow in the appear-

ance of a very distant point when the eye is continuously lowered. At *1* a single erect image would be seen, but when the eye falls to *2*, a second image will suddenly appear far below it; further progress will be attended with the rise of the lower image, which of course is an inverted one, until it unites with the undeviated one and both vanish simultaneously. It is not difficult to extend the explanation to the case where the eye is at rest and the distant object has appreciable dimensions. Suppose, for example, that the distant object is a ship and that the eye is in the region *1*; then the upper parts of the masts and sails would, if the ship is tall enough, be seen single; but as the observer directs attention to the lower portions he would notice that at a certain level an inverted repetition appears some distance below, and that the intervening space is occupied with the erect image above and the inverted below which comes up to meet it much as it would appear to do were the ship resting upon a horizontal mirror, except that there would be a region of indistinct vision between the two.[1]

From what is said in the last paragraph about the physical condition which attends the production of such a layer arrangement of the air, it is quite evident that these particular phenomena can be seen only over extensive plane surfaces which are kept continuously at a temperature much higher than the adjacent air. This, however, can be also brought about when such a flat surface is continuously heated by the sun, and nothing is more common to those who look for them than such effects over smooth pavements or along sun-heated walls. But there is another instance which has been known for an indefinite time and often described — that of the desert mirage. Over extensive plains thus heated by

[1] This last statement will become evident to the mathematician who reflects on the character of the locus of the centres of the curve *bc*; it is a curve having its vertex near *o*, but running out asymptotically to both of the right lines *oc''* and *ob''*. Moreover, it is easy to see that this image in the neighborhood of the *b''b* will be strongly astigmatic, whence comes a very obvious vertical streakiness here. Often this is the first appearance of the approach of conditions favorable for observing the more complex phenomena.

the sun the rarefied air adjacent to the ground gives an inverted image of the sky near the horizon, and, if there happens to be an object beyond the sensible horizon which raises itself through this lowest stratum, it, also, will appear to be reflected, and thus in a surprising degree of likeness imitate the effect of a distant and quiescent sheet of water. A moderate wind enhances the effect in general, because it serves to keep the upper air at a uniform temperature, while it will be so checked close to the earth by friction that it does little toward cooling the air in contact with the ground; of course a high wind is unfavorable.

Turning now to the case where the lower strata of air are denser than those aloft, we are brought to the consideration of more complicated, more interesting, but, at least in our latitudes, much rarer phenomena. Here, of course, there is no mechanical unstability, since the denser layers are everywhere below the less dense; thus we shall not be surprised to learn that these appearances are far steadier, and afford much better optical images. It is well to note that the preceding phenomena as well as those which follow can be best, and sometimes only, seen by aid of a telescope of moderate power. Even an opera glass is an efficient aid.

Figure 33 represents this case diagrammatically. Here an eye in the region *1* will see a single undeviated image of the distant object, but at *2* the beginning of a confused image in the elevated position indicated by the direction of the line $c''c$. This latter will quickly resolve itself into an erect image above and an inverted image below, which at *3* will be about midway between the two upright images. As the eye rises, the inverted image will approach the lowest one and join it at *4*. Above this position there will be but a single image slightly flattened in a vertical plane; finally, above *6* all evidence of extraordinary refraction will disappear. The condition of affairs between *4* and *6* is worthy of a moment's attention. Since the path of this portion of the waves is concave toward the earth, sometimes strongly so, an object below the true horizon may be seen, as is generally

true of a rising sun or moon, and many mariners report having seen and recognized ships under such circumstances. The extraordinary flattening of the rising full moon, especially in hot autumn evenings, is often noticed. It is also the explanation of the famous Fata Morgana.

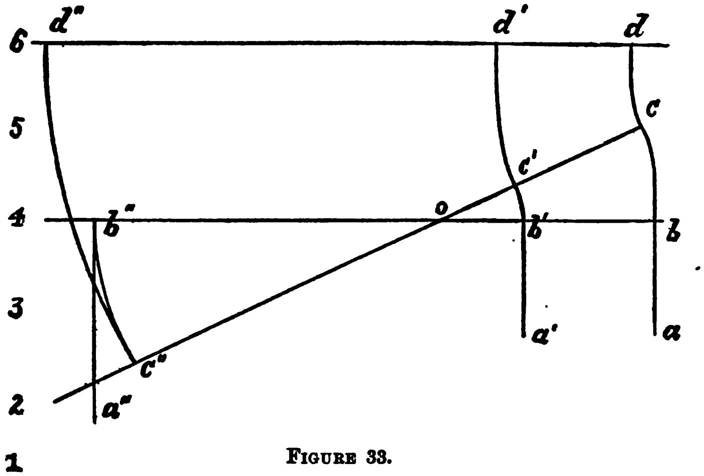

FIGURE 33.

The favorable meteorological conditions for the production of this class of phenomena consist in the existence of air at a high temperature over a cold sea; they are therefore not often seen in our southern waters, but over the colder waters of northern bays and those surrounding England they are less infrequent. It is also evident that any circulation of the air attending a wind would be fatal to an extensive development of the appearances; hence one can hope to see them only during calm weather. Thus it is that the most surprising accounts of like phenomena come to us from arctic explorers who have had opportunities of observing over a partly frozen sea in calm and warm days of spring or summer. Among such observers, Scoresby was the first who published a carefully written account of his observations, which have since served as the basis of theoretical discussions by numerous writers. In view of what has been previously said concerning the scale

of the figures, it is of interest to note that this observer never found well-developed cases when the object was less than ten to fifteen miles distant, which makes the distance of the region of the air that acts like our tank about half as much. Almost all the cases which he pictures are in perfect agreement with the conclusions deduced from Figure 33, but there two cases of apparent disagreement which are worthy of further consideration. The first is represented by a number of figures in which the erect image appears below, with an inverted image above, the second erect image which should be seen surmounting all being absent. This, however, is readily explained if we imagine the line *dd''*, which marks the lower limit of the region of uniform density, to be lowered, or, what amounts to the same thing, the part *cd* of the wave-surface to be much shortened; in this case the image in question would be much diminished in height and brightness, and might well escape detection if the observer were not guided by theoretical knowledge. The other case is that of four images in a vertical line. This may be explained by the assumption of the existence of another transition stratum above, an assumption which does not appear at all improbable, in view of the fact that such an arrangement would be in stable equilibrium. Indeed, it requires considerable care to avoid the production of just such extra-transition strata in the syrup-water tank; while the presence of two enables the experimenter to copy the case of four images very satisfactorily.

The explanation of the singular phenomenon known as scintillation, or twinkling of the stars, has given a vast deal of trouble to philosophers of both ancient and modern times. In its most ordinary manifestations it appears as a continuous but extremely irregular variation in the apparent brightness of the stars when not too near the zenith. This appearance is generally absent in the case of the planets, even when strongly marked in the fixed stars, but it may be seen sometimes in terrestrial sources of light, provided that they are of very minute angular dimensions and sufficiently re-

mote. When the phenomenon is quite regular, a bright star appears to undergo continuous change of color with only moderate changes of brightness. If an observer presses one eye slightly with the finger so as to see two images of a twinkling star, or attains the same result by placing a thin prism before one of the eyes, it will be seen that the changes in one of the images are quite independent of those in the other. Finally, in telescopic images of scintillating stars, at least in telescopes of moderate aperture, quite similar phenomena are visible even in faint stars. These may be varied in an instructive way by shaking the telescope slightly so that the image of the star is stretched out into a ribbon of light, in which case the curve described by the image appears to be made up of most vividly colored elements. For a long time it was supposed that Arago had given a thoroughly satisfactory explanation of the whole series of phenomena by attributing them to interference of light waves which enter different portions of the pupil or objective, and which must have experienced different retardations in their passage through the atmosphere; indeed, for more than two generations this explanation was regarded as one of the triumphs of the undulatory theory of light. Respighi first demonstrated the inadequacy of Arago's theory, by a spectroscopic examination of scintillating stars, the results of which he embodied in a long series of conclusions of which the most important are the following: —

I. In spectra of stars near the horizon we may observe dark or bright bands, transverse or perpendicular to the length of the spectrum, which travel more or less quickly from the red to the violet or from the violet to the red, or oscillate from one to the other color; and this, however the spectrum, may be directed from the horizontal to the vertical.

II. In normal atmospheric conditions the motion of the bands proceeds regularly from red to violet for stars in the west, and from violet to red for stars in the east; while in the neighborhood of the meridian the movement is usually oscillatory or even limited to one part of the spectrum.

From these statements it is at once evident that the rotation of the earth is concerned in the phenomenon, and, although the previously accepted explanation does not necessarily exclude this as a part of the cause, the fact that it did not suggest it must be taken as indicating a lack of completeness if nothing more. But the interesting and unexpected character of these observations led Lord Rayleigh to study critically the physics of the problem, and to recognize that Arago's premises are quite untenable, and his theory therefore wholly fallacious. The true explanation is to be looked for in the optical irregularities of the atmosphere, combined with the obvious fact that the paths of the short wavelengths which reach the eye are, on account of the not inconsiderable dispersive power of air, rather widely separated from those followed by the longer wavelengths. In every case the path of the more refrangible light lies above that of the less refrangible. If, therefore, there is a limited volume of air which alters the form of a wave-front passing through it so that less of that particular light enters the eye at a stated moment, a very short time later the rotation of the earth will have brought this air to a higher point if it lies toward the west and to a lower if toward the east. Thus, in accordance with Respighi's observations, the spectrum of a star in the east would show progressive changes from violet to red, while in an opposite quarter of the heavens the effect would be reversed. The whole series of changes would occupy but a short time, in general less than a second. The orderly change in the spectrum of a star on the meridian would not follow, since in this case the motion of the earth does not carry the disturbing body of air from the lower level of the paths of the longer wavelengths to a higher one, where it would modify the more refrangible light. Since a very small angular displacement of a star would entirely alter the position of the light-paths with respect to a small body of air, it is easy to see why different points of the disk of a planet may scintillate in so perfectly unre-

lated a way that the light from the planet as a whole does not change in intensity.[1]

Of bodies suspended in the atmosphere which may be the source of important optical phenomena, drops of water and crystals of ice only are of sufficiently frequent occurrence in our climate to be of especial interest. To the first we owe coronas and rainbows; to the second, halos and their accompanying appendages. We may direct our attention to them in the order here named.

Coronas consist of one or more colored circles concentrically surrounding the sun or moon when these are covered by light and transparent clouds. They are readily distinguished from the larger circles which are called halos, not alone on account of their smaller and variable sizes, but also by the fact that the inner edge is blue and the outer red, an order which is reversed in halos. Fraunhofer showed that these coronas may be exactly imitated by scattering very small, circular, opaque bodies, such as lycopodium powder, in a perfectly irregular manner over the surface of glass and looking through such a plate at the sun or moon. By employing a telescope so as to magnify the effect, and by choosing a much smaller source of light — a star or planet, for example — he was able to secure the same result when the powder was replaced by a large number of equal disks of tin foil scattered in a perfectly irregular manner in front of the objective. These experiments demonstrate at once that we have to do with an effect of diffraction, and we must therefore look to the laws which govern this class of phenomena for a rational explanation. It is necessary to add little to what appears in Chapter III. to make the matter clear.

In that chapter it was shown that a bright point seen through a very small round hole would appear as a disk surrounded by a series of rings, blue on the inner edges and red on the outer. If another hole of the same size were perforated in the screen, we found that the disk and rings remain,

[1] Further considerations with reference to the phenomenon of scintillation may be found in Appendix B.

but are doubled in brightness and crossed by a series of dark lines. If the number of holes is increased, taking care to keep them of the same size, the disk with its concentric rings remains, but with greatly increased brightness; while the system of dark lines has gradually become so complicated and condensed that it ultimately becomes indistinguishable. Thus the final effect would be the same as that of a single hole of the size chosen, but with the brightness multiplied by the number of apertures. This conclusion, the validity of which has been mathematically established by Verdet, may be readily tested by pricking a large number of holes in a bit of tin foil with the point of a fine sewing-needle, taking care to secure equality of size by thrusting the needle through the same distance each time, and then looking through this screen at an artificial star. A similar effect may be produced by looking at a bright star with a telescope through a screen of paper perforated with irregularly distributed holes even as large as a tenth of an inch in diameter; if the planet Jupiter is selected as the object, the resulting image is an almost perfect imitation of a well-developed lunar corona.

To make use of the facts just established in the explanation of coronas, it is necessary to state and explain a fertile method in this department of optics, known as the principle of Babinet. This principle may be thus given: If, on account of the presence of an opaque screen, however complicated, we have light at any point between the source and that point, then, if all the transparent portions of the screen are made opaque and the opaque portions transparent, the quantity of light at the point will remain unchanged. The proof of this fact becomes obvious if we consider Figure 14, page 32. The reason why light is found at p_1 is only because the wave-surface is limited; consequently the opaque portion of the screen cuts off just what would, if added to the disturbances which reach this point, exactly destroy them; in other words, the waves suppressed by the screen are of like intensity to those which reach the point in question, but differ from them by a half wavelength, or at least by an odd num-

ber of half wavelengths, in phase. Thus, if the interchange between the opaque and the transparent parts of the screen is made, the intensity of illumination at these points remains unchanged. Although this is but a special case, the character of the reasoning is perfectly general, and we are therefore led to accept the principle as completely established and applicable to all cases.

By means of the principle of Babinet we pass at once from the case of the screen irregularly pierced with uniform circular air holes to the glass plate covered with irregularly dispersed disks or spheres of uniform size, and finally, to small spheres of water suspended in the atmosphere. That the spherules of water are essentially opaque follows at once from the fact that only an infinitesimal part of the light which passes through them can reach the eye, on account of their short focal length. The necessary conditions for the existence of a corona are, therefore, suspended particles of water of very uniform size, so small that the diffraction rings produced by them shall be considerably larger in angular dimensions than the sun or moon. As the spherules grow larger the corona becomes smaller, and *vice versa;* thus we have in the coronas a guide as to whether precipitated moisture is increasing or diminishing, and an indication of value in predicting the changes of weather.

There are other cases of phenomena of the kind under discussion and bearing a similar interpretation. Most observers would probably recognize a system of colored rings surrounding an electric arc-light seen against a dark background, especially immediately after sleep. Often after an eye has suffered some injury — as from a blow, for example — these rings appear with unaccustomed brightness, and remain for a long time with decreasing intensity and concurrent increasing dimensions, until they resume their normal appearance. They are attributed to slight opacity in the epithelial cells of the cornea. Occasionally one may observe such circles about a light when seen through a sheet of glass upon which there is a considerable deposit of moisture from the air,

although this indicates a regularity of the deposit, which is rather unusual.

Rainbows are produced by refraction and reflection of direct sunlight falling upon spherical drops of water. Ordinarily very little light reaches the eye from such an illuminated drop, for that which leaves the drop, either after simple reflection from the first surface or after more complicated refractions and reflections involving the rear portion, will be so widely scattered that a very minute fraction of the whole can fall on the small area of the pupil. There are certain definite directions, however, where this conclusion

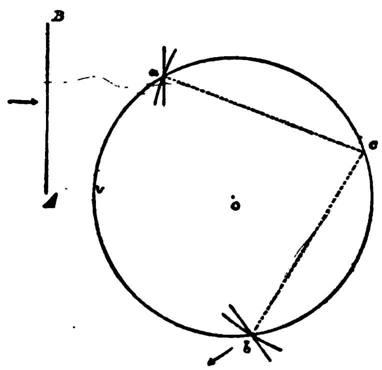

FIGURE 34.

does not hold. The accompanying Figure 34 and Figure 35 will help to make this statement clear. In the former let *o* be the centre of a drop of water, and *AB* a plane wave-surface moving from left to right and ultimately falling upon it; in entering the sphere every portion will have its curvature modified and its direction of propagation altered. It is not difficult to recognize that these modifications will increase with the distance of the portion of the wave-surface from the vertex *v*. Admitting the validity of this conclusion, there must be a portion of the wave-surface, such as that represented at *a*, for example, in which the curvature produced by refraction will be just sufficient to

make its centre fall exactly on a point at the rear surface of the drop, whence a portion suffers reflection toward the front of the drop. It is this reflected portion alone which further concerns us. This, as follows at once from relations of symmetry, will reach a point *b*, at the same distance from the point at which reflection takes place — *c* in the figure — as is *a;* here it is submitted to a second refraction, which just restores the portion of the wave-surface to a plane. From this point the light will be propagated without further change of shape, as far as the vertical direction is concerned; and to an eye placed anywhere in its course the drop will appear to be a bright object. Light which falls on almost any other point of the drop will be so spread out after reflection that it will add very little to that due to the former source. Calculation shows that for any transparent sphere of a substance possessing the optical constants of water the points *a*, *c*, *b*, are so related that the total change of direction of motion of the light is nearly 138°. Generalizing from these facts, we are led to the statement that every drop upon which the sun is shining and which happens to be 138° from the sun will appear to emit light, and as all such drops lie in a cone whose apex is at the eye, the result will be a circle of light at the angular distance given, or, what amounts to exactly the same thing, a circle having its centre directly opposite and with a radius of 42°. An obvious consequence of the theory is that no rainbow of this kind can be seen when the sun is more than 42° above the horizon.

Before passing to the explanation of another kind of rainbow, a remark on the brightness of the bow may be made. Although it has been pointed out that the light which forms the rainbow is propagated without loss in the meridian plane, on account of its flatness in that direction, there is no such conservation in a plane perpendicular to this. Hence one might conclude that the intensity of illumination in a bow would rapidly fall off with increasing distance of the drops which produce it, nor would there be any escape from this conclusion were there not the compensatory fact that with

increasing distance the number of drops which are near enough to the required direction to take part in the phenomena would increase in nearly the same ratio.

The case considered is not the only one in which a plane incident wave-surface may emerge as a plane in an altered direction, as will appear from Figure 35. Here, at a larger distance from the vertex v, a greater deviation and change of curvature is produced by refraction, so that the wave-surface is concave until it reaches a point c_1, after which it is convex until incident on the surface of the sphere where it suffers another change of curvature by a partial reflection. If the proper geometrical conditions are met, this second

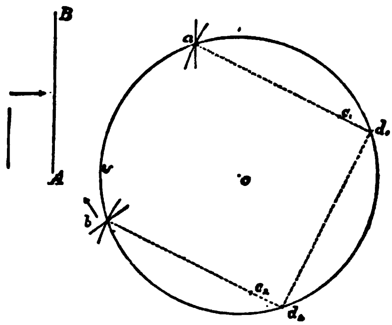

FIGURE 35.

change in curvature — always in the direction of a diminished value, since the surface is concave — may be just sufficient to restore the wave to its original state of flatness, in which case it will follow the course indicated by the dotted line and, after experiencing a reflection at d_2, which would again reduce the curvature by a like amount so that the centre is at c_2, will attain the surface at b, where a final modification would restore the original flatness. The number of different ways in which this same general result may be accomplished is boundless, the geometrical conditions being that, if there is an odd number of interior reflections, the middle one must correspond to a point where the curvature of the wave-surface is infinite; while, if the number of interior reflections

is even, the middle chord which marks the path of the light corresponds to a region where the curvature is zero. The two preceding figures are constructed to represent the cases of one and two interior reflections for water. It will be observed that in the first case the direction of the emergent wave is changed more than a right angle from its original direction, but less than two right angles; in the second case the deflection is between two and three right angles. Calculation yields the following values of the deflections for a number of reflections varying from one to five: —

No. of Reflections.	Deviations.
1	138°
2	231°
3	318°
4	404°
5	486°

These are given to the nearest degree for red light.

From this table we may conclude that there is a second locus of drops which send light to the eye, namely, a circle everywhere 231° from the sun, or, what is exactly equivalent, at an angular distance of — 51° from the point opposite to the sun. The minus sign attaching to the last number means that the light which comes to us from the second bow has entered the drops on the lower side, instead of on the upper as in the first case. The second rainbow is necessarily much fainter than the first, or inner one, since its light has experienced two partial reflections instead of one; still, it is always visible when the primary is very bright.

What is true relative to the faintness of the secondary bow is true in a far higher degree of those bows of a higher order, the existence of which is demonstrated by the foregoing table, and this is quite sufficient to account for the fifth bow never having been detected; but were the third, and perhaps even the fourth, as favorably situated for observation as are the

primary and secondary, it would certainly be visible on occasions. When one notes, however, that both these are little more than 40° from the sun, and that this portion of the sky must be very strongly illuminated by the light which has passed through the drops, it is easy to see that the conditions are most unfavorable.

This explanation of the rainbow we owe in nearly the same form to Descartes; but as he understood nothing of the laws of dispersion, it was left for his great follower, Sir Isaac Newton, to bring the theory of the rainbow nearly to completion. The color of the bow is obviously due to the differing refrangibility of different wavelengths, and it is quite easy to extend the foregoing reasoning to include the consequences of this fact. For example, in the figures given to illustrate the paths of the light in raindrops, and which are drawn for red light, it is evident that since the change of direction by refraction is always greater for short than for long waves, in the first one, the deflection for blue light in the primary bow would be more than 138° (almost 140° in fact) and its radius would be less than 42°; hence the primary bow would have the prismatic series of colors, beginning with red on the outside and terminating with violet on the inside. For the secondary bow, in the same way, we would have a deviation greater than 231° (about 234°) for the blue, and consequently a circle with a radius of more than 51°; hence the secondary bow is red on the inner side and violet on the outer. With the added statement that the prismatic colors cannot be very pure on account of the considerable angular diameter of the sun, we have explained, with one exception, all the prominent and familiar features of the rainbow. Before describing this and certain other less striking phenomena, it will be necessary to consider another point which the diagrams teach.

In certain determinate directions the intensity of the light leaving the drops is much greater than in all others; but this is equivalent to saying that near certain points all por-

tions of the incident wave are deflected by nearly the same amount, or, in other words, that the angle of deviation changes very slowly with the angle of incidence. This, however, is the characteristic of maximum and minimum values of a variable, as, for instance, when the sun is changing its distance from the horizon most slowly it is either at its highest or lowest point; or when the days are changing in length most slowly, they are either at their longest or shortest. In the case under consideration the deviations cannot be maximum values in either the one or the other, for a single reflection of that portion of the wave which passes through the centre of the sphere would give a deviation of 180°, and after two reflections of 360°. These values being respectively greater than 138° and 231°, we conclude that the particular deviations which are efficacious in producing the bows are minimum values. From this we may make the highly important deduction that no drops nearer to the sun than 138° can send any light to the eye by means of a single reflection, and none nearer than 231° can send any light by means of two interior reflections; consequently those drops which are situated between the two bows, being nearer in one direction than the first limit and in the other than the second, can send no light either by a single or a double interior reflection. This is the explanation of the relative darkness of the sky between the bows, which forms one of the most conspicuous features of a well-developed rainbow.

Since the light which produces the bows comes from portions of the plane wave-surfaces which have suffered a minimum deviation, it follows that there are portions of these surfaces above and below the most efficacious part which have the same deviation, such light reaching the eye from drops just inside the primary and outside the secondary bow. But such portions will have passed through slightly different lengths of water, and thus have experienced slightly different retardations. Such conditions give rise to the phenomena of interference. It is not difficult to see that this will produce, inside the primary and outside the secondary, a series of

repetitions of each color, of rapidly diminishing brightness, and at an angular distance decreasing with increasing size of the drops. Repetitions of this character are called supernumerary bows, and they are not infrequently seen under the highest part of a bright inner bow and more rarely above the vertex of the secondary bow. The conditions of distinctness are smallness and uniformity in size of the drops, which conditions are more likely to be found well up in the air. Another effect of interference is slightly to change the dimensions of the bows, decreasing the apparent diameter of the inner and increasing that of the outer. The changes, although small for ordinary rainbows, may amount to several degrees for the white bow, which will engage our attention next.

A lunar rainbow appears almost colorless on account of its faintness, just as foliage loses nearly every trace of color by moonlight, but this is a merely physiological effect. Occasionally, however, a very bright primary bow is seen with only a tinge of red on the outer and of blue on the inner side. This occurs when a bright sun shines on a dense wall of fog where the droplets are sufficiently large to give a tolerably definite reflection, but at the same time differ greatly among themselves in size. In accordance with the fact stated in the preceding paragraph such a bow is always smaller in angular diameter than a colored bow.

Sometimes arcs of rainbows are seen on a sward when covered with dew, and also in the spray thrown up by the bow of a vessel. In such cases, although the images formed on the retina are always circular arcs, we are apt to ascribe to them the form of their projections upon the surface below; hence, if the sun has an altitude of more than 42°, they appear to be arcs of ellipses; if less than this, arcs of hyperbolas.

The theory of the supernumerary bows, which completed the work of Descartes and of Newton, was first suggested by Young in 1804 and completely worked out by Airy in 1836.

The phenomena caused by crystals of ice suspended in the

air are much more complicated than those due to spheres of water, and some of them are of even more frequent occurrence, at least in high latitudes. All these phenomena, complex as they are, may be divided into two classes, namely, those produced by crystals whose axes are directed purely fortuitously, and those due to crystals which for some reason have their axes parallel to a fixed line or to a fixed plane. The former of these two classes may be explained in a few words, since the theory is very simple.

Crystals of ice assume a very large variety of forms, as is well known to every one who has observed the shapes of snowflakes, but they are all built upon a general plan which has the right hexagonal prism for its foundation. It will be necessary, however, to assume the presence of the simplest form only in the following discussion, that is, of simple hexagonal prisms; but of these we shall suppose that there are

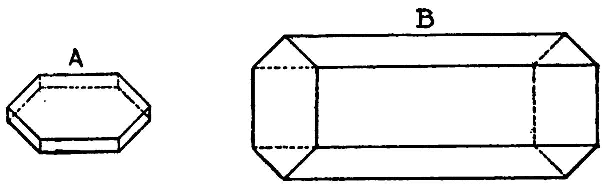

FIGURE 36.

two types such as are represented in Figure 36, where *A* is a flat hexagonal plate and *B* a rod-like or elongated prism with a hexagonal base. Both these forms have been repeatedly observed. If one of these transparent bodies is considered as an optical apparatus, we perceive that it presents six sixty-degree prisms (in the sense in which this term is used in optics), and twelve right-angled prisms, the former kind being limited by a lateral face and its alternate and the latter kind by a lateral face and either base. If flat wave-surfaces fall upon any one of these eight faces, provided that the angle of incidence lies within a certain determinable range of values, it will emerge from the other face of the prism deviated by

an amount which admits of ready calculation, since it depends upon the angle and upon the refractive power of ice. Thus for a sixty-degree prism we find a minimum deviation of 22°, corresponding to an angle of incidence at the first surface of about 41°, since the index of refraction of ice is 1.31. If this angle of incidence becomes either greater or less, the proportion of light transmitted — nearly total in the first case — very rapidly diminishes. If, therefore, there should happen to be such a crystal of ice suspended in the air at an angular distance of 22° from the direction of the sun, provided that it happened to have its face in the proper direction, it would send all the light which enters at one face toward the eye of the observer. At a very small distance further from the sun the crystals might send a much diminished quantity of light to the eye, but at a still less distance none whatever, since 22° is the minimum deflection. If a host of such crystals were uniformly distributed through the air, it is obvious that we should have a circle of light 44° in diameter, with the sun at its centre. When we take into consideration the secondary phenomenon of dispersion, we conclude that this ring must have a red inner border and a much less distinct bluish outer border. This constitutes the familiar halo about the sun or moon which may be seen perhaps as often as fifty to a hundred days a year in our latitudes. The explanation of the phenomenon is due to Mariotte.

The rectangular prisms which are formed by the dihedral angles at the bases of the crystals give rise in a strictly similar way to a ring with a radius of 44°, since the low refractive power of ice permits light to pass through a prism even as obtuse as this.[1] On account of the far more critical arrangement required for a ninety-degree prism to be effective, as well as on account of the much greater area of the sky occupied by this second halo, it is very rarely seen in its entirety; but certain portions of it, as will appear later, are often conspicuous. The first suggestion of this explanation is attributed to Cavendish. These two concentric rings are

[1] See, however, remarks in Appendix C with regard to this point.

the only features which are due to ice crystals whose axes have purely fortuitous directions.

It is worth noting that, as no crystals nearer the sun than 22° can send light to the eye by refraction, the portion of the sky immediately inside the inner ring is darker than the region exterior to it. This is a very conspicuous feature in a well-developed halo.

Hexagonal plates of the shape of *A* in Figure 36, and elongated prisms like those represented at *B* in the same figure, would not, in general, fall through the air with their major dimensions indifferently directed with respect to their line of motion. If the air were quite calm the plates would fall with their axes vertical, that is to say, the plates themselves would be horizontal, while the second form would fall with their axes horizontal. Of the latter type we may even add that the prevailing arrangement may be that of a maximum section remaining continuously horizontal, and with it, of course, two of the lateral faces of the prisms, although in many cases — doubtless the majority of cases — we must imagine the crystals constantly rotating on their horizontal axes in their downward motion. The reasons for these definite arrangements are readily explicable, though we need not stop to give them here. It is sufficient to recall the fact that such a body, as a match for example, cannot be made to fall through any considerable distance endwise, nor can a small flat body, like the petal of a flower, for example, be made to fall edgewise.

The phenomena resulting from the refraction of light by such directed crystals must obviously bear some relation to the directions involved, which are only that of the sun and the vertical direction, or that of the zenith; hence all the resulting forms must be symmetrical with respect to a plane through these two points on the celestial sphere. This geometrical consideration will lighten the description of the more complex features of halos.

As the vertical height of the sun above the horizon must have a material influence on the aspect of the halo, we are

obliged to consider the problems presented when this element is varied, in order to secure a quite definite notion of what may be seen. In turning our attention first to the halo of a low sun, we shall meet not only with the features which are most frequently recorded, but we may gain much by a preliminary study of some well-described halo of complete development. None is better suited to our purpose than a very extraordinary one seen by Parry and Sabine in 1820, at their winter quarters at Melville Island, and described by them with a scientific precision hardly found elsewhere. Figure 37 is a

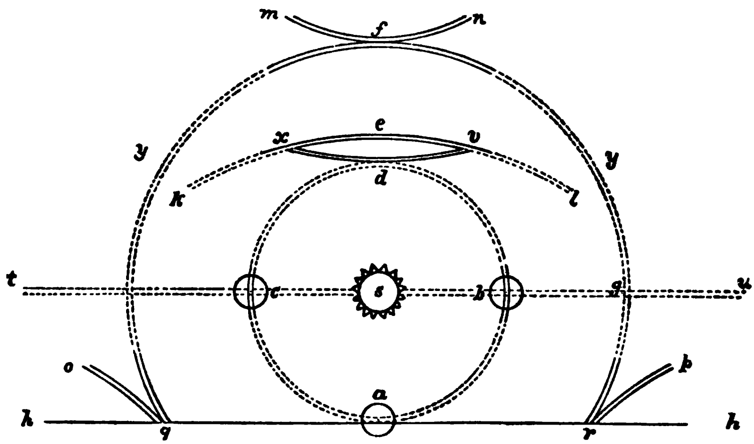

FIGURE 37.

faithful copy of their drawing, and it is followed by a transcript of the accompanying description:[1]—

"From half-past six till eight A.M., on the 9th, a halo, with parhelia, was observed about the sun, similar in every respect to those described on the 5th. At one P.M. these phenomena reappeared, together with several others of the same nature, which, with Captain Sabine's assistance, I have endeavored to delineate in the annexed figure.

s, the sun, its altitude being about 23°. *h*, *h*, the horizon.

[1] Parry's First Voyage, pp. 164–165.

t, *u*, a complete horizontal circle of white light passing through the sun.

a, a very bright and dazzling parhelion, not prismatic.

b, *c*, prismatic parhelia at the intersection of a circle *a*, *b*, *d*, *c*, whose radius was 22½° with the horizontal circle *t*, *u*.

x, *d*, *v*, an arch of an inverted circle, having its centre apparently about the zenith. This arch was very strongly tinted with the prismatic colors.

k, *e*, *l*, an arch apparently elliptical rather than circular, *e* being distant from the sun 26°; the part included between *x* and *v* was prismatic, the rest white. The space included between the two prismatic arches, *xevd*, was made extremely brilliant by the reflection of the sun's rays, from innumerable minute spiculæ of snow floating in the atmosphere.

qfr, a circle having a radius from the sun, of 45°, strongly prismatic about the points *fqr*, and faintly so all round.

mn, a small arch of an inverted circle, strongly prismatic, and having its centre apparently in the zenith.

rp, *qo*, arches of large circles, very strongly prismatic, which could only be traced to *p* and *o*; but on that part of the horizontal circle *tu*, which was directly opposite to the sun, there appeared a confused white light, which had occasionally the appearance of being caused by the intersection of large arches coinciding with a prolongation of *rp* and *qo*.

The above phenomena continued during the greater part of the afternoon; but at six P.M., the distance between *d* and *e* in creased considerably, and what before appeared an arch, *x*, *d*, *v*, now assumed the appearance given in fig. 12, plate 287, of Brewster's Encyclopædia, resembling horns, and so described in the article 'Halo,' of that work. At 90° from the sun, on each side of it, and at an altitude of 30° to 50°, there now appeared also a very faint arch of white light, which sometimes seemed to form a part of the circles *qo*, *rp*; and sometimes we thought they turned the opposite way. In the outer large circle we now observed two opposite and corresponding spots *y*, *y*, more strongly prismatic than the rest, and the inverted arch *m*, *f*, *n*, was now much longer than before, and resembled a beautiful rainbow."

In order to make clear the following discussion of the phenomena due to directed crystals, we must find a means of

distinguishing the different faces. This may readily be done by reference to the accompanying Figure 38, where p_0 to p_5, inclusive, represent the six lateral faces of the hexagonal prisms, alternate faces having either even or odd suffixes; o represents a base, and if we wish to indicate the opposite base which is the only remaining face of the prism, we shall designate it by o'.

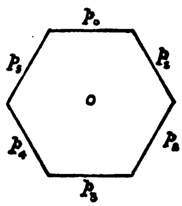

FIGURE 38.

It is not necessary to dwell upon the crystals of the first group, that is, those which have a purely fortuitous arrangement of their axes and the presence of which is clearly indicated by the concentric circles *acdb* and *qfr*, since these phenomena have been already considered; but we may turn at once to the more complicated features. As the simpler of the others, we will review the effects due to the second group, that is, to plate crystals figured in *A*, Figure 36, which fall with their axes perpendicular. Of all such crystals which happen to be at a distance of 22° to the right or left of the sun, a very great number would be found at a given instant in such a position that light entering one vertical lateral face would emerge from an alternate face with a total deviation of 22°. These produce the prismatic parhelia *c* and *b*. Crystals of the same kind slightly more distant from the sun might send light to the observer in two different ways, namely, either by refraction at an angle differing from that of minimum deviation, or by refraction at the entering face followed by total reflection on the adjacent face and a final refraction at the face of emergence. In the former case light of less refrangibility would be superimposed upon that from prisms more nearly oriented to the position of minimum deviation; consequently light from these directions would tend toward colorlessness. In the latter case the light would be pure white; since the dispersion produced at one refraction would be compensated at the other. From these considerations we see that if this group of crystals were

very well represented the parhelia ought to have a vivid prismatic red color on the side toward the sun and a bright colorless tail stretching out for a number of degrees in the opposite direction. Since the last-named feature was not noted by the observers, although it is most conspicuous in a large number of recorded halos, we must conclude that the group *A* was not very well represented in the case under discussion. Each of the vertical faces of these crystals would somewhere produce by reflection an image of the sun which must lie at a height above the horizon equal to that of the sun itself; hence the total effect would be a colorless horizontal circle passing through the sun and extending completely around the heavens. This circle, called the parhelic circle, is represented in part in the drawing by the band *tu*. Finally, we have only to add that the effect produced by the rectangular prisms, one side of which in this case must remain constantly horizontal, is a colored arc *mfn*, which, since these prisms run to an edge and are not interrupted by an intermediate face capable of adding totally reflected light, is far purer in its colors than are any of the other portions of the halo, often rivalling the brightest rainbow. The reason this light is spread into an arc instead of being confined to a bright spot similar to the twenty-two-degree parhelia is because in this case the edges of the prisms are only restricted to parallelism to a horizontal plane, not, as in the other case, to parallelism to a fixed line. It is interesting to note that sometimes this arc and the parhelia attributed to the same crystals constitute the whole visible phenomenon, indicating that only this particular group is present.

We now turn to the consideration of the effects produced by prisms of the third group *B*. We shall assume that some of these fall with their maximum sections continuously horizontal, in which case we have the two bases constantly vertical and two lateral faces, which for convenience we will indicate as p_0 and p_3, constantly horizontal. Such crystals as happened to be at the point indicated by *e* in the drawing,

and at the same time had their axes nearly at right angles to the vertical plane of the sun, were the source of the brilliant prismatic image there observed, which was more remote from the sun than *d*, because these crystals were not at the position of minimum deviation. Crystals of the same kind to the right and left of these, which happened to have for the moment a proper orientation of their horizontal axes, produced similar prismatic images at a somewhat greater distance from the sun but at a less altitude, whence the arc *xev*. All the light which entered a lateral face, was reflected interiorly from a base, and then emerged from a face parallel to the first, came from some point in the parhelic circle, as did also such light as was reflected from a base without entering the prism. It is obvious that much of the light which suffered this modification was totally reflected from the interior and would yield very bright images; indeed, it is probable that in this particular halo almost all the parhelic circle is to be attributed to this group. So, too, most of the light in the arc *mfn*, produced by rectangular prisms with horizontal edges, therefore also identical in action with the corresponding portion of the crystals of group *A*, must be attributed in this particular halo to *B* crystals. Finally, the arcs *qo* and *rp* of the drawing were produced by the action of this group by means of the rectangular prisms which were underneath and had their edges at a constant angle of 60° with the horizon. These are all the features pictured in the drawing which owe their origin to the *B* group, although we shall return to the discussion of another one mentioned in the text, which under certain circumstances becomes of great importance.

The stability of the *B* type of crystals with respect to the horizontality of the maximum cross-section can be hardly regarded as very pronounced because of the small excess of this section over the minimum, which is but 30° from it; therefore we are not surprised to find that many ice crystals having this shape do not fall in the manner indicated. Whatever the conduct of these may be, unless, in-

deed, they simply oscillate about the position of stability, the effect produced by them will be exactly the same as though they rotated at a constant speed around their horizontal axis. To specify this sub-type of columnar crystals they will be designated by the symbol B'. Their existence in this particular halo is not only demonstrated by the phenomena attributable to them alone, but they were in certain favorable situations clearly seen as "innumerable minute spiculæ of snow." To these were due the brighter portions of the forty-six-degree halo in the neighborhood of its tangent arcs, the arc *xdv* tangent to the twenty-two-degree halo, and the brilliant white light above the last of which the colorless spot on the horizon below the sun is the correlative. The white light was due to light totally reflected from the interior of crystals which happened to be favorably situated, exactly in the same way as explained in the case of the A group immediately outside the parhelia, those sufficiently near the observer sparkling in the sunshine. The anthelion, a bright spot in the parhelic circle directly opposite the sun, was also produced by these B' prisms. The manner of production may be explained as follows: Imagine a B' crystal anywhere in the line from the observer to the anthelion, which for the instant has two of its faces, p_0 and p_8, for example, perpendicular; light entering one of these faces would suffer successive reflection, in either order, at the opposite face and a base, and then emerge through the face of entry. Such light would come from the anthelion. Now imagine the crystal to rotate on its horizontal axis; it will obviously recover an exactly equivalent position after turning 60°, but in the mean time the image formed by the two refractions and the two intermediate reflections would describe an oblique path passing through the anthelion. Since this reasoning is independent of the direction of the horizontal axes with respect to the vertical circle through the sun, all this kind of crystals in the direction indicated would give rise to oblique arcs having the point horizontally opposite the sun in common; hence this region would appear brighter than the surround-

ing sky. Since they lie within certain limits in all possible directions, these oblique arcs would not in general be distinguishable, but with a low sun it is not difficult to prove that those having an inclination of about 60° with the horizon are brighter than the others, and such arcs seem to have been recorded a considerable number of times, although they must be classed as among the rarer appendages of halos.[1]

Now consider the changes recorded by Parry and Sabine as the sun sank toward the horizon. First of all, *e* separated more widely from *d*, which is quite in accordance with the theory, because the sun is still further removed from the position of minimum deviation, corresponding to an altitude of the sun of about 49°, than in the case represented by the drawing. Then the points indicated by *yy* in the outer circle began to grow bright; this was the beginning of a phenomenon complementary to the tangent arcs *qo* and *rp* produced by a refraction through a lateral face in momentarily favorable position and a base. A third feature de-

[1] As we shall not have occasion to return to this point, and as it has given no little trouble to investigators, it may be worth while slightly to extend the reasoning. It is obvious that if the two interior reflections were both partial the emergent light would be extremely faint and probably quite negligible; therefore the interesting cases are when one of these reflections is total. Since, however, there can be no question of total reflection on a lateral face, because light which can enter such a face can evidently emerge from the opposite one, we are confined to cases of total reflection on a base. The minimum angle of incidence in these prisms under consideration which will yield such total reflection is 57°.7; hence only those whose axes have a direction in azmuth differing from that of the sun from 32°.3 to 90° meet the condition defined. From those anywhere near the latter limit, however, very little light would be returned on account of the small effective area of the base, which is only the projected area; hence much the brightest arcs must be ascribed to those having an angular displacement from the sun's vertical not less nor very much greater than 32°.3. With a higher sun it is possible to have light other than that which has entered a vertical lateral face fall upon the interior rectangular mirror, namely, some of that which has entered at the superior adjacent face p_1, and, after reflection, emerged at p_5. In such cases it would be possible to have two pairs of oblique arcs through the anthelion, which also seem to have been observed on rare occasions. Finally, we may add that an anthelion with a high sun, say of 50° or more, is impossible, since there will be either no light at all or an altogether insignificant amount incident upon the two faces.

scribed in these observations is of especial interest because so conspicuous when B crystals are abundant and the sun very high, namely, the faint arcs which seem to stretch toward the anthelion in the direction of a prolongation of the tangent arcs to the lower part of the forty-six-degree halo. These were produced by light which entered the B crystals through a p_0 face and, after an intermediate reflection from one of the bases, emerged through p_2 or p_4. Such arcs would be very faint throughout, but brightest at a distance of about 90° from the sun.

The long discussion of this complicated halo which we have given serves two different purposes: First, it gives the explanation of a considerable part of all the recorded phenomena of halos; second, it justifies our assumption of the three types of crystals, and two varieties of the last type. We may therefore with much confidence approach the explanations of the remaining phenomena, although there are perhaps no complex cases on record nearly as well observed as the present one.

As the sun attains a greater altitude above the horizon, a larger and larger portion of the light entering an upper horizontal face is reflected from an adjacent vertical face; hence, in general, the parhelic circle gains in brightness with this change, especially as the interior reflections become total when the altitude is more than 32°. In the presence of A crystals some of the light which enters an upper base may suffer successive reflections from two adjacent sides before emerging from the inferior base. This gives rise to a singular phenomenon not infrequently recorded, which is known under the name of the paranthelion of 120°. To explain it a well-known law of optics must be stated. If an image of an object is formed by two reflections from two plane mirrors, its angular displacement from the object, measured in a plane perpendicular to the line of intersection of the mirrors, is exactly twice the angle included between the normals to the mirrors. The phenomenon discussed in connection with Figure 3, page 9, is a

special case where the angle between the mirrors is a right angle. In the case in question the angle between two adjacent sides is 60°; hence the images formed of the sun would lie in the parhelic circle 120° from the sun or 60° from the anthelion on either side. It is the latter relation that gives them their name. Another feature which very unusually developed and abundant *A* crystals might produce is a pair of bright spots at *g* and its corresponding point on the opposite side of the sun, which are called the parhelia of 46°. These are probably secondary images of the inner parhelia, that is to say, images of them formed by refraction exactly in the same way as they are formed with the sun as the source of light.

With a sun much higher, say half-way toward the zenith, the series of phenomena will be very considerably modified. The *A* crystals still yield their parhelic circle with its parhelia well outside the twenty-two-degree circle, also the parhelia of 120°; but the tangent arc of the forty-six-degree circle can no longer appear. The appendages produced by the *B* group are also rather inconspicuous, although they add to the light of the parhelic circle and also produce short oblique arcs through the anthelion. The crystals which have been designated by *B′* may still give rise to an anthelion as before; but their most notable contribution to all the phenomena is an oval touching the twenty-two-degree halo at its highest and lowest points, and lying wholly outside of it. This oval, which has been very frequently seen, has been perfectly explained by Venturi, who showed that the upper (*kxdvl* of Figure 37) and lower tangent arcs of the twenty-two-degree circle, that have been shown as attributable to the horizontal crystals without fixed direction of the axial sections, bend toward each other with increasing altitude of the sun, until at a height of about 50° they unite and form the oval. It is easy for us to see that some such phenomenon must follow, if we reflect that, were the sun at the zenith, all the *B′* crystals, like the undirected crystals, would serve simply to produce a twenty-two-degree halo; but as the sun

moves from the zenith those crystals which are in the same vertical circle would alone maintain the same relation to it, those not in this vertical transmitting the light somewhat obliquely, and consequently giving a greater deviation with a resulting locus of the image of the sun at a greater distance than 22°. Hence, from considerations of symmetry, we recognize that this locus must be an oval.

At a still higher altitude — 60° or more, for example — the *A* crystals are of little significance except as they may add to the effect produced by the rectangular edges of the *B* group and to the intensity of the parhelic circle. Paranthelia are invisible or inconspicuous on account of the thinness of these prisms and the large angle of interior incidence. On the other hand, the *B* group of crystals, if largely represented, becomes very important. First, they may produce a strong parhelic circle; second, they yield from light which enters a base and emerges at a lower horizontal face a tangent arc, brightest with a solar altitude of 68°, to the lowest part of the forty-six-degree circle; finally, from light which enters at p_1 or p_5, is reflected at a base, and emerges at the lower horizontal face, they produce a pair of peculiar spiral curves, which intersect at the anthelion point, but fade out before reaching a second crossing of the vertical circle passing through the sun. Figure 39 represents a spherical projection of a carefully constructed halo of this type, the constants of which were calculated from the present theory. The zenith distance of the sun is 25°; about the sun as a (projected) centre are drawn the unbroken circles representing the twenty-two-degree halo, while with the zenith as a centre and a radius of 25° the double circle which represents the parhelic circle is described in broken lines. The remaining curve, self intersecting at S', gives the two spirals under discussion. The parts drawn in smooth lines are the brightest portions; most of the remainder, drawn in broken lines, is relatively faint, although in the region immediately below the sun it coincides with an arc which is produced by incident and emergent refractions without the

intermediate reflection. Were the oval of Venturi also present we should here find three distinct arcs which are the

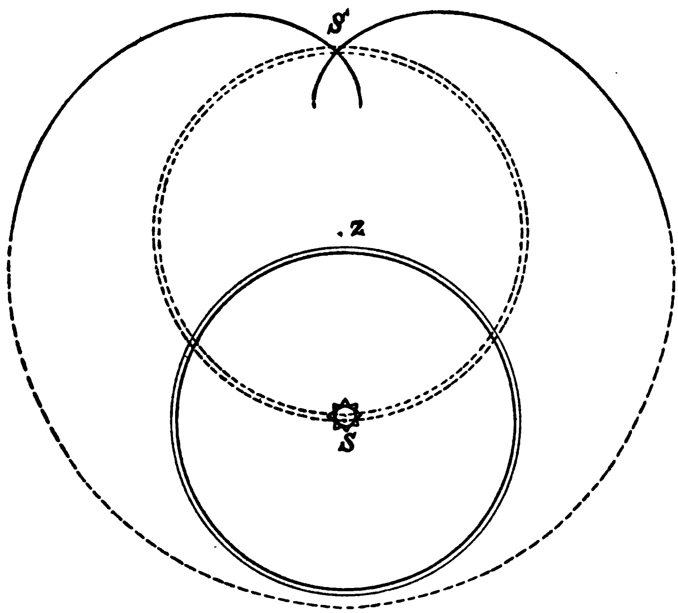

FIGURE 39.

exact complement to the three arcs in the region *de* of Parry and Sabine's drawing.[1]

[1] Halos of this kind are practically invisible in Europe, on account of the high latitude of the countries where alone these phenomena are seen, but in the United States, where the climate is like that of northern Europe and the latitude that of southern Europe, many such have been visible. Sketches of extraordinary halos of this type may be found by the curious reader in the "American Journal of Sciences" (1), vols. vii., x., xi., and (2) vol. xl., p. 394; and in the "Report of the British Association," 1861 (2), p. 63. None of these, unfortunately, is very critical, since in every case the observer seems to have assumed as a matter of course that all the arcs seen must be parts of true circles, and in not a few cases he has not hesitated to complete the circles in accordance with his own conceptions. Nevertheless the fact that the sole feature in which all these sketches are in agreement is that derived from the theory developed above, and pictured in the drawing, may be taken as almost conclusive evidence in its favor. Further remarks on the theories of halos will be found in Appendix C.

There remains a feature of undoubted authenticity, that is, one which has been recorded frequently by competent observers, which is worthy of note, namely, vertical columns of light occasionally passing through the sun when not far from the horizon. These may be explained as follows: Since the larger surfaces of the *A* group of crystals are continuously nearly horizontal, if there happen to be many of them near the observer and between him and the sun, they would yield a specular image of the sun which would appear above or below that body according as it happens to be at the moment below or above the true horizon. If the crystals possess a slight rocking motion in their fall — and we have abundant evidence that this sometimes occurs, not only by such features as the brightening of the forty-six-degree halo, as explained on page 146, but also by the arcs known as the tangent arcs of Lowitz, which extend from the parhelia accompanying a high sun downward to the twenty-two-degree halo — the specular images will be stretched out into a vertical column in much the same way as an image of the sun produced by a surface of quiet water is altered into a streak of light whose axis passes through the sun when this surface is ruffled by ripples; and this quite irrespective of the direction of the motion of the ripples. It is evident that the reflection from the lower surfaces of the prisms in question will be total; hence we must attribute to them the chief rôle in this phenomenon, and very little to the faces of the *B* group, some of which are also horizontal.

CHAPTER VIII

THE EYE AND VISION

In its general structure the eye is a camera obscura, but it possesses an important difference from the common camera in that the interior is filled with media not unlike water in their optical properties. This peculiarity carries with it certain interesting modifications in the phenomena of vision, which, however, can hardly be more than barely indicated here. If we look at Figure 40, which represents a horizontal section of the human eye, copied after Helmholtz's careful drawing, we shall be enabled to understand the principal details. The whole body is a spheroid somewhat flattened in the direction of its axis. Within is found an anterior chamber *B*, which is filled with a liquid little different from pure water, and hence called the aqueous humor; and the posterior chamber containing a somewhat firm transparent body to which the name vitreous humor is given. The latter is marked *C* in the diagram. Between these two is situated a transparent lenticular body *A*, which is known as the lens. The front of *B* is bounded by a very transparent convex skin called the cornea, which projects quite perceptibly beyond the general surface of the eyeball. Resting upon the surface of the lens, and constituting the back of the anterior chamber *B*, is a delicate membrane *bb*, having a circular hole in its centre. The central aperture is called the pupil of the eye, and the membrane which it pierces, the iris. It is the latter which gives the characteristic color to the eye, and, as a part of the optical apparatus, its office is to vary conveniently the effective diameter of the lens. The

variation in the size of the pupil is brought about by the contraction of one or the other of two sets of muscles, the first being radially disposed so that contraction enlarges the aperture, while the other, arranged in a circular manner along the inner edge of the iris, causes a diminution of the pupilar opening by contraction. These muscles respond auto-

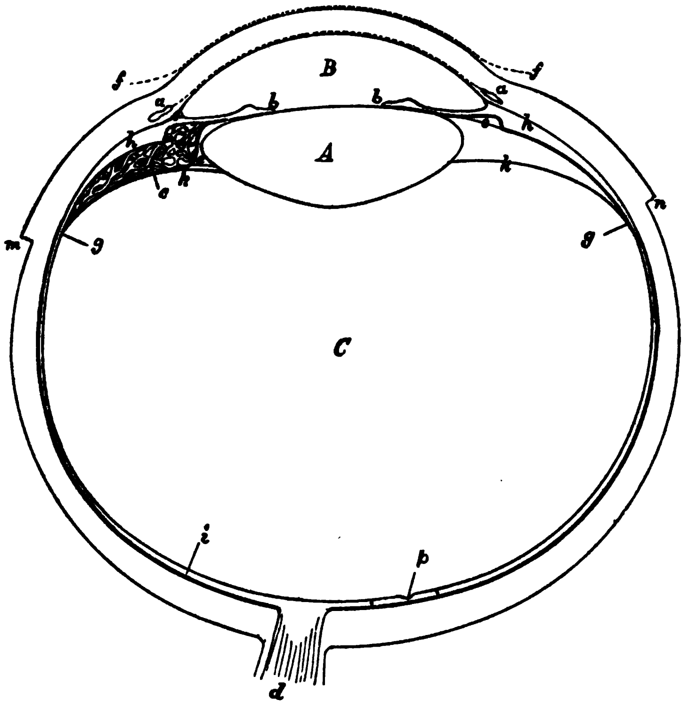

FIGURE 40.

matically to the stimulus of light on the eye, so that when this is very bright the pupil becomes extremely small. The ordinary range of effective diameter of the pupil may be taken as from somewhat less than a twentieth of an inch in bright light to a little more than four times as much in a very faint light, thus increasing the illumination more than sixteen fold.

The cornea and lens combine as an optical system which alters the curvature of nearly flat wave-surfaces falling upon them, so that after modification their geometrical centres fall just on the back inner surface of the eye, and consequently form there an extended image of objects sufficiently remote. Over this rear surface is spread a delicate membrane *i*, composed chiefly of nerve tissue which is connected with the brain by a bundle of nerve fibres passing through the orifice *d*. The membrane is called the retina, and the bundle of nerves the optic nerve. The aperture for the optic nerve lies on the nasal side of the axis of the eye; hence the figure represents a horizontal section through the right eye when seen from above. Before describing more in detail the construction and function of the retina, we may consider the provision for securing sharpness of definition in the images on the retina, when the distance of the object viewed is changed. If the cornea and lens together possess just sufficient power to cause flat wave-surfaces to have their centres on the retina, they would prove insufficient for convex wave-surfaces such as would come from nearer objects. The means taken in a photographic camera to adjust for such differences is to alter the distance between the lens system and the screen which receives the images, but in the eye this necessary adjustment, called accommodation in this case, is wholly different and without analogy in any artificial optical instrument. In short, accommodation is produced by a change in the absolute power of the lens attending a variation in its thickness. Figure 41 clearly shows this change, the left half exhibiting the shape of the lens when suited to distinct vision of a distant object, and the right half the modified shape for near objects. It will be noticed that the alteration is almost confined to the anterior portion of the lens. The capacity for adjusting the power of the lens to the immediate demands upon it is very remarkable in young children, enabling them to see objects with perfect sharpness from a distance of three or four inches up to an indefinitely great distance; but throughout life it suffers a diminution

which seems to follow a fairly constant rate. If during the earlier period of life the eye is used almost exclusively for near objects, the lens is apt to assume a permanently thickened form, so that it is too powerful for waves having a remote source. Such an eye is called near-sighted, or myopic,

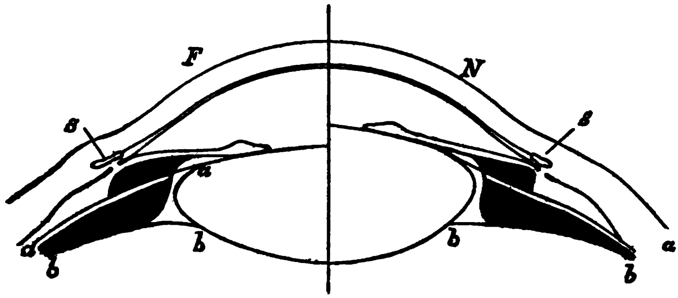

FIGURE 41.

and in order to see distant objects well its power must be reduced by means of a negative lens, that is, of a spectacle glass thinner in the middle than at the edge. This is the common origin of the near-sightedness which frequently makes its appearance during youth among those of studious or sedentary habits. If this predisposing cause of myopia is wanting, the gradual loss of the power of accommodation with advancing age results in the adjustment becoming more and more close than proper for remote objects. An eye with this sort of limitation is called far-sighted, or presbyopic. To adapt it to clear vision of near objects its power must be increased by the aid of a positive lens, that is, one which is thicker in the middle than at the edge. The change will generally have progressed so far by the fifth decade of one's life that the nearest point of distinct vision is at arm's length, when the necessity of aids to vision becomes first evident even to those who are very unobservant. One or the other of the foregoing descriptions is the common history of the normal eye, but many persons are born with myopic eyes, generally on account of abnormal axial length of the eye-

ball; likewise, instances are not rare where a defect of the opposite kind, too short an axis, is congenital. In the latter case positive lenses will be required for seeing even remote objects clearly. We may note, in passing, that another defect, unfortunately far from uncommon, is that in which the eye has different powers in different planes. This can be recognized by the differing distinctness with which horizontal and vertical lines may be seen in a brick wall, or by the elongated aspect of a bright star. The defect is known as astigmatism, and it may be corrected by the use of a spectacle glass bounded by cylindrical surfaces.

The rear portion of the inner eye is immediately surrounded by a thin, black, membranous lining between the retina and the firm coat (sclerotic coat) which forms the body of the eye, the office of which is to shield the retina from all light not entering through the pupil. This dark membrane is named the choroid. It is evident that the pupil appears quite black to an observer because of this membrane as well as because of the fact that the only points of the retina visible to such an observer are those covered by the image of his own pupil. The first condition, namely, the presence of the choroid, is wanting in albinos, and the second, a sharp retinal image, in animals whose vision is imperfect or who have their eyes adjusted for something widely remote from the spectator. In these cases the pupil no longer seems black but of the pink color of the retina behind, and a singular glare appears which is sometimes thought to have its origin in the eye itself. It is easy to devise means of illuminating the retina by placing either a transparent mirror between the eye of the observer and the observed eye, or a mirror with a small perforation which permits the observer to look through it. Instruments for this end are called ophthalmoscopes, and are of great importance to the surgeon.

The blood-vessels which serve to nourish the retina lie on its anterior surface, while the portions that are the true sensory organs are confined to the posterior surface. For this

reason light which enters the eye from any point produces shadows of these vessels in the retina. These are usually unrecognized, both from their ordinarily ill-defined nature on account of the angular dimensions of the pupil, and because the portions of the retina immediately behind these vessels, consequently those most frequently affected, have adjusted themselves to the relative darkening due to their presence. If, however, one looks at the bright sky through a pin-hole in a card, or if an image of a bright object at the extreme margin of the field of vision sends its light down upon the retina, the shadows become surprisingly obvious, especially if the source of light is rendered intermittent, or, in the former experiment, if the card is kept in constant motion. An admirable way to secure the effect is by means of a lens to form an image of a brilliant light on the ball of the eye, just back of the edge of the cornea; the light which penetrates the sclerotic coat will produce wonderfully distinct shadows.

Just opposite the centre of the pupil at the back of the eye there is a slightly depressed portion (*b*, Figure 40) called the yellow spot, or macula fovea. This is the region of most accurate vision, and visual impressions derived from images not restricted to this spot are always somewhat ill defined. Our knowledge of the color has been derived primarily from dissection of the eye, but an interesting evidence of its nature may be found by putting a piece of blue glass over the pin-hole in the perforated-card experiment, when it will be found that this spot, which lies at the precise centre of the field of vision, is indicated not only by the relative freedom from blood-vessels, but also by its distinctly darker shade. The latter distinction is due to the greater absorption of blue light by this portion of the retina, thus demonstrating its yellow color.

The region at which the optic nerve enters the back of the eye is called the white spot, or macula lutea. This spot, occupying a very considerable area in the visual field, is wholly insensitive to light, and consequently marks a blind

spot in the field. That every human eye possesses an area which is quite blind and is large enough to render invisible an area seventeen times as large as that covered by the full moon, or as large as would be occupied by the image of the head of a man at a distance of six or seven feet, would doubtless surprise most people not informed in this department of science. In the year 1668 Mariotte, while endeavoring to find whether the optic nerve proper is sensitive to light — a scientific question which seems to have occurred to no one before him — made this unexpected discovery, which excited a very lively popular interest at the time. It is easy to demonstrate the existence and position of this spot by means of a card having a mark and an appropriate object,

FIGURE 42.

as illustrated in Figure 42. If in this the right eye is fixed upon the cross, the left eye being closed, and the distance of the figure from the eye is varied, it is easy to find a position at which the large white circle entirely disappears, although one can still see the regions surrounding it. A few persons can recognize the existence and position of the blind spot, without the aid of a definite figure, by careful attention to the visual field when one eye is closed.

Experiment shows that the seat of visual sensation is restricted to an extremely thin layer at the back of the retina; therefore the ultimate structure of the nerve tissue in this region is of especial interest. A powerful microscope shows that it is composed of two types of elements, which are repre-

sented in Figure 43, one of which (a) is called a rod and the other (b) is called a cone. We are to understand that the whole sensitive portion of the retina consists of rods and cones radially disposed with respect to the centre of the eye; but it is found that within the yellow spot only cones are present, while these elements become relatively few as one approaches the peripheral portions of the retina. For this reason, together with the fact that the eye is found to be more sensitive to feeble illumination in regions outside the yellow spot, it is assumed that the cones have to do with the minute perception of form, and the rods with perception of differences of illumination. This inference is strengthened by the observed fact that the angular distances of adjacent cones correspond fairly accurately with the minimum angular distance at which two objects can be seen other than single. This distance, one of the most important constants of the eye, is about one minute of arc, as has been already pointed out on page 95.

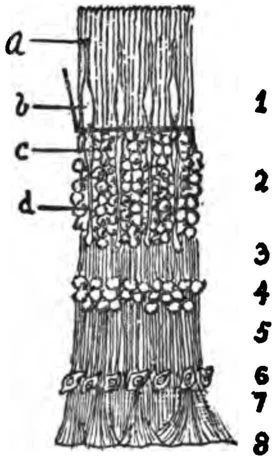

FIGURE 43.

Many visual phenomena depend on the fact that the sensitiveness of the retina falls off with remarkable rapidity after the beginning of excitation. For example, after remaining in darkness for a long time the light of the full moon produces a painfully strong impression, but under other circumstances the light of a bright noonday may be less impressive although 620,000 times as intense. A retinal impression lasts for a perceptible time after the cessation of the stimulus, however, as appears from the fact that a rapidly moving bright point seems to carry with it a streak of light; and if the path is a closed curve in which the point travels as frequently as ten or more times a second, the streak seems to be also a

closed curve of nearly uniform brightness. Such experiments prove that the sensation endures without very great diminution for a time as long as a tenth of a second. Images which appear upon the retina after the cessation of the light producing them are called after-images. Very intense after-images often last many seconds, but they undergo a succession of changes in color and intensity that are highly complex, and which cannot be regarded as belonging to strictly normal vision. Such images are distinguished from the former kind by the term negative after-images, and although their study has proved of value in several fields of physiological optics, they will not be further considered here.

There is, strictly speaking, no such thing as an optical centre to the eye, but there are two points separated by an interval of about a sixtieth of an inch, so related that light which would pass through the first before refraction appears to pass through the second without change of direction after refraction. These points are called the first and second nodal points, respectively. A point half-way between them may be taken in most instances as the optical centre of the eye, so that light having a direction which would carry it through this point will remain unchanged in direction after its final refraction. Since this centre is well in front of the geometrical centre of the eye about which it rotates, a distant object just hidden behind a screen close to the eye becomes visible by oblique vision when the eye is turned away from its direction, even when the head remains fixed.

As there is no provision for eliminating the secondary phenomenon of dispersion which accompanies all cases of refraction, the eye is not achromatic. Although ordinarily overlooked, it is not at all difficult to devise experiments which render this defect conspicuous. For example, should half the pupil be covered by an opaque screen, the image of a linear object seen against a bright field — a window bar against the sky, for example — if parallel to the edge of the screen, will appear blue on one side and yellow on the other. But the obvious effects due to this imperfection in

the eye are noted when observing a surface having figures of bright red color upon a bright green or blue field; in this case a peculiarly disagreeable sensation arises from unsuccessful efforts to accommodate for both colors at the same time.

A structural defect in the eyes gives rise to certain visual phenomena even more conspicuous than the chromatic errors just noted, that is, a lack of homogeneity in the lens itself. This body, built up of transparent cells in a highly complex manner, has a roughly stellate structure, which, even in the best eyes, is the immediate cause of the irregularly pointed aspect of bright stars. No one of these defects seriously impairs vision, provided that the accommodation is perfect; but when this, too, is at fault, the light from a point, instead of being uniformly distributed over a small portion of the retina, is somewhat irregularly gathered in areas radially placed in accordance with the lens structure. It is for this reason that the horns of a crescent moon appear multiple unless the eye is perfectly adjusted as regards accommodation. Other defects, which vary largely with different individuals, are found in the presence of opaque bodies within the media of the eye, which, although they cast shadows upon the retina, are ordinarily overlooked on account of the quickly acquired insensitiveness under continuous impression; still, they can always be detected if the eye is turned suddenly from the contemplation of a dark field to a bright sky. These evanescent objects are called muscæ volitantes.

If in a darkened room the eyes are closed and the eyeball is slightly compressed with the finger-tip at a point as far removed from the pupil as possible, a deep violet-colored spot surrounded by a ring of bright light is recognized in that portion of the visual field which corresponds to the part of the retina thus disturbed. Again, if in a dark room the eyes are turned suddenly as far as possible, either to the right or left, similar spots will be seen which are produced by a strain brought upon the retina by the stretching of the

optic nerves. These observations are interesting illustrations of the important physiological law that an excitation of a sense organ results in some kind of sensation proper to that organ, and is quite independent of the nature of the stimulus. In the cases described, of course, there can be no possible analogy between the ordinary stimulus produced by light and this mechanical disturbance; so, too, singular light sensations, and light sensations only, result when an electric current is directed through the eye. The subjective phenomenon known as "seeing stars" is produced by a disturbance of the elements of the retina attending a violent shock, although it may be perhaps due to a sort of reflex action from the brain.

The study of color sensations offers a field of great interest, in which Sir Isaac Newton made the earliest and most important investigations. He demonstrated that light of different refrangibilities produces different color sensations, and that white light can be decomposed by a prism so as to give rise to an infinite number of color sensations. To seven of these he attached familiar color names. He also showed that these seven colors combined in proper proportions would cause the sensation of white light. He even went further than this, and proved that if the spectral series of colors is expanded by the addition of purple, it is always possible, having chosen one color, to find a second which, combined with it, would produce white light. Pairs of colors so related are called complementary colors. Although Newton's contribution to the theory of color sensation is by no means restricted to the discoveries named, we shall find it more convenient to consider the general theory developed in modern times.

Critical study of color sensations has led modern writers on chromatics to revise Newton's terminology of the spectral colors in replacing his seven by the series red, orange, yellow, green, blue-green, blue, violet. This rejects Newton's indigo and supplies blue-green, or robin's egg color, which is not only quite as properly an independent color as is

orange, but is very convenient because exactly complementary to vermilion red. Since Newton had proved that with two complementary colors it is possible to secure by combination, not only either one of the colors, but also any gradation between it and white, and that with three colors it is therefore possible to produce a far greater range of colors, the question naturally arises as to the minimum number of primary colors by the use of which all possible colors might be imitated; or, to state the problem in a much more scientific form — as color is, after all, a question of sensation only — what is the smallest number of primary color sensations that by appropriate combination will yield all possible color sensations? Dr. Thomas Young seems first to have stated this problem clearly and to have given its solution. His assumptions are the following: —

I. The eye possesses three kinds of nerve termini. Excitation of the first gives rise to the sensation of red, of the second to the sensation of green, and of the third to the sensation of violet.

II. Monochromatic light, that is, light of a single wavelength, stimulates these three kinds of nerve termini in differing ratios, with different lengths of the waves. The nerves sensitive to red will be most strongly affected by light of greater wavelength, the green-sensitive nerves by light of medium length of waves, and the last by the shortest waves. It is not assumed, however, that any spectral color fails to excite one or more of the nerve elements; on the contrary, in order to explain a vast number of phenomena, it is necessary to assume that every spectral color excites simultaneously all the elements, but to a varying degree.

To render clear the second of the fundamental assumptions we may refer to Figure 44, which is copied from Helmholtz. In it the letters at the bottom represent the spectral colors, and the curves above them the relative sensitiveness of the three hypothetical nerve elements; whence we deduce that simple red stimulates the red-sensitive nerves strongly and the others feebly, with a resulting sensation of red, and

simple yellow excites moderately strongly the red and green elements and feebly the violet, with a sensation of yellow as a result. Other colors may be discussed in a similar way. It is proper to remark that the figure carries the term blue instead of violet in the lowest curve, but this is not material in the theory, nor is it yet determined which is preferable. We shall hereafter employ the convenient although unscientific terms, red nerves, green nerves, etc., to designate these elements, which are at the foundation of the theory.

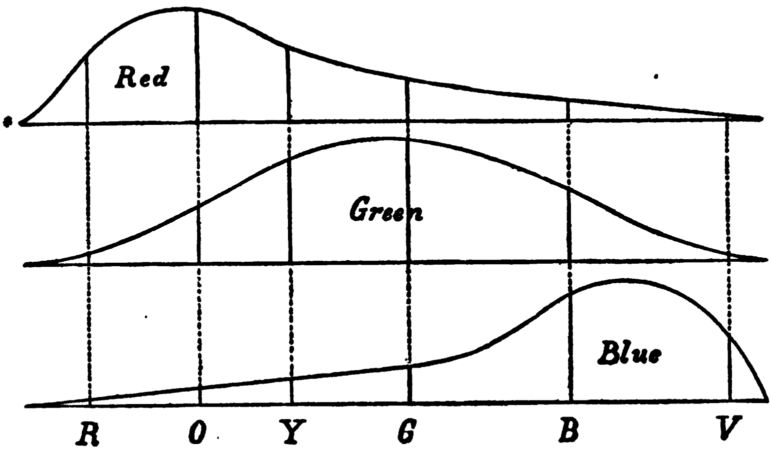

FIGURE 44.

This is a sketch of the famous theory known to physicists and physiologists as the Young-Helmholtz Theory of Color Sensation. On account of its celebrity, its simplicity, and its general agreement with experience, we shall describe it somewhat at length, although it is quite certain that it must be regarded only as a first approximation to the explanation of a highly complex series of phenomena.

An obvious corollary from the theory is that the normal eye never has a given fundamental color sensation without at the same time having in a less degree a combination of the other two. For example, vivid green light falling upon the

retina awakens a moderate sensation of a mixture of red and blue with the stronger green, which we may regard as combining with a portion of the green sensation to produce white; hence the sum of the sensations may be characterized as a pure green mixed with a certain amount of white. It follows from this that were it possible to stimulate the green nerves alone the resulting sensation of greenness would be far more intense. In effect this result can be readily produced by taking advantage of the rapid loss of sensitiveness of the retina under stimulus. If one looks for a few minutes at a sheet of magenta paper, which is a nearly equal combination of red and of blue, the eye will have lost a considerable part of its sensitiveness to these two colors, so that if green light falls upon the modified retina, the sensation of greenness will be astonishingly enhanced. The same principle of course holds true for all pairs of complementary colors, that is, either one of such a pair would appear with increased strength of color after the retina has been slightly wearied by contemplating the other.

Before applying the foregoing theory to the explanation of colors we must acquaint ourselves with certain useful terms employed in the discussion of chromatics, and with some of the experimental methods of testing the conclusions deduced from the theory. It has been found that all color sensations can be defined by the use of three terms only: The first is hue, that is to say, that which gives the ordinary color name; the second is saturation, or freedom from admixture of white light; the third is luminosity, which is the equivalent of brightness when applied to white light. A few illustrations derived from common color names will render clear the meanings of these important terms. A straw-color, for example, has the hue yellow, the saturation very low, and the luminosity very high; olive-green has a greenish yellow hue of high saturation and low luminosity; pink is a rose-red hue of low saturation and high luminosity. These examples are perhaps enough to serve our present purpose, but many others will appear later.

We may now turn to a consideration of some of the methods employed in experimenting with these mixed color sensations, for it will be observed that the whole theory rests upon the assumption that the retina is stimulated by the simultaneous action of unlike colored lights — a very different thing from the action of light coming from a mixture of unlike pigments. Thus we shall see presently that the combination of yellow and blue lights may produce a sensation of pure whiteness — would be, in short, white light — but it is a fact known to every child that light from a mixture of yellow and blue paints is in general green. The reason for this singular difference, which so often leads to a quite false conception of the requirements of a sound color theory, will appear when we come to consider the cause of color in natural objects. Any experimental tests of the theory, therefore, must be made by an actual mixture of the colored lights in question. This may be evidently done by reflecting light from one source on to a screen which is illuminated by light from a second source at the same time, and almost as simply in a number of other ways; but none is so convenient and of such universal applicability as the method invented by Maxwell. The principle at the basis of this method is the persistence of visual impressions. Imagine a disk of cardboard half of which is covered with yellow paper and the other half with blue paper; if this is rotated about its centre many times a second, it would appear to the eye uniformly tinted with yellow of half the luminosity possessed by the yellow paper added to blue of half the luminosity which would be presented by a disk entirely covered with the blue paper, and the combination would be found free from all color, that is, it would be white. But this white would clearly be of a low luminosity; in short, in comparison with white paper, it would be called gray. If we have two cardboard disks, each possessing a radial slit and a uniform color, the two may be put together so that either one will exhibit a sector of the circle as large as desired. This composite disk rotated rapidly about its centre will enable us to get the result of

the combination of the two colors in any required and easily determined ratio. The defect of the method lies in the fact that all such compound colors possess a luminosity less than the brightest of the components, but the convenience is so great that it offers the best-known method of exhibiting the principal phenomena. Small disks of colored papers not more than an inch in diameter dropped axially upon a spinning top are admirably adapted for the experiments to be described.

The choice of the three fundamental colors is to a certain extent arbitrary — the only condition imposed is that it shall be possible to produce white by a proper combination involving all three — but Maxwell's choice of vermilion-red, emerald-green, and ultramarine-blue, all of which may be readily procured, has much to recommend it; and various phenomena presented by color-blind eyes render it practically certain that these are not far out of the way. If, according to the method described, red and blue are combined in varying proportions, starting with the whole of the red disk showing and ending with this completely covered with the blue one, it will be found that we have a series of hues passing continuously from pure red through rose-red, magenta, and violet, to pure blue; if the red disk is replaced by a green one and a similar experiment is made, we pass from blue through the hues greenish blue, blue-green, bluish green, to pure green; finally, if the blue disk is replaced by the red one and a similar progressive change is made, we pass through a series of colors in the order yellowish green, greenish yellow, yellow, orange, to a pure red. Should all three disks be employed at once, it will be found quite easy to arrange them so that the combined impression is a pure gray; in other words, a white of low luminosity.

This series of experiments teaches us that it is possible to secure all hues by a properly chosen combination of the three fundamental colors, but it is very far from presenting us examples closely approximating many of the most familiar colors. We have, it is true, secured all possible hues, but

only a small part of all possible colors. The distinction lies in the fact that we control by this means only the relative proportions of colored lights, whereas the other two elements defining a color sensation are not varied at will; hence this general review of colors is very restricted. It may, however,

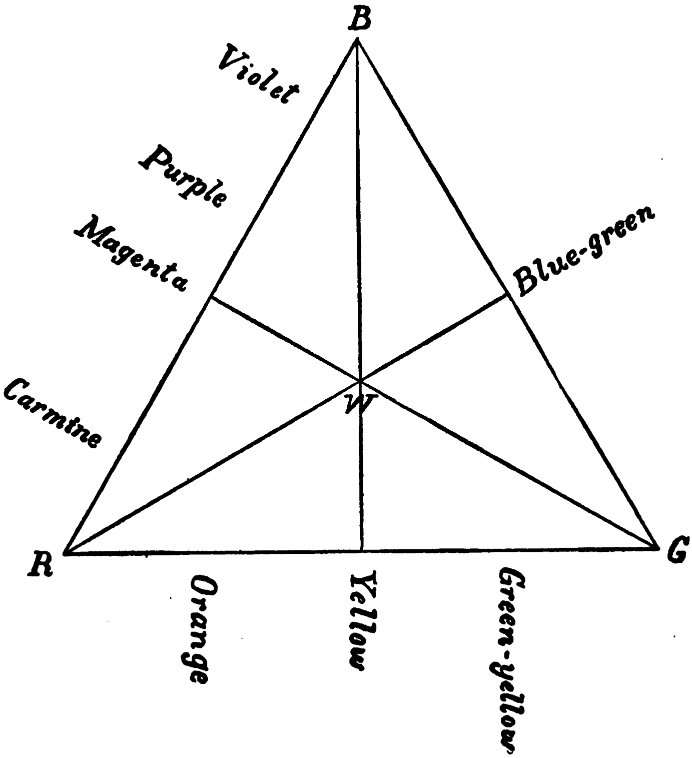

FIGURE 45.

be greatly extended by the use of a diagram invented for this purpose. Suppose we place the three fundamental colors at the vertices of an equilateral triangle, as represented in Figure 45, in which these colors are designated by their initial letters, and select the centre as the point to indicate white; then, if we make the assumption that any point on

one of the sides of the triangle represents a hue produced by mixing the two adjacent primary colors in an inverse ratio to the distance of the point from either of the corresponding apexes, we have an absolutely unambiguous method of defining all possible hues. If in addition to this we let, conventionally, the linear distance from the central point represent the saturation, or freedom from admixture of white, every point in the plane of the triangle will represent a color of which two of the three necessary constants are given in a perfectly definite manner. A number attached to a given point in this diagram may be used to designate the luminosity. From such a diagram much may be learned. For example, it gives at once all pairs of complementary colors, since any straight line drawn through the central point cuts the sides of the triangle in two points which define colors thus related. Suppose one demands the color complementary to emerald-green; the diagram shows it to be a color on a line passing through G and W and on the opposite side of the latter point; all this region represents a hue which is an equal mixture of red and blue, called magenta when highly saturated and lilac when pale. The region from green to blue-green would find its complementary hues between magenta and pure red, while those on the other side of the green would be complementary to hues lying between magenta and ultramarine-blue, that is to say, to the purples and violets. A line drawn from B through W intersects the RG line near its middle point, the region of yellow. It is just this portion of the diagram which appears to be least satisfactory as a means of representing familiar colors, and it is therefore worthy of more careful attention. According to the theory an equal mixture of red and green should form a yellow, but an experiment after the Maxwell's disks method gives a dingy brownish yellow color of which the relation to the familiar bright yellow so extremely common in flowers is by no means obvious; and it is quite impossible to secure a good yellow by this means. This is because the luminosity of yellow is very great, whereas light from green paper is

much less bright and that from red paper is far more inferior. There is no difficulty in demonstrating the identity of hue of the mixed color and the bright yellow, however, by a somewhat indirect method. If we take three smaller disks, one white, one black, and the third yellow, and superimpose them upon the larger disks of red and green, it is always possible to adjust the former so that in rapid rotation the inner disk matches the outer circle. Figure 46 illustrates this interesting experiment, which proves that the yellow produced by the combination of the fundamental red and green is less saturated than that of the bright yellow, since it was found necessary to add white to the latter; again, the mixed color was much less luminous because it was found necessary to reduce the brightness of the yellow paper by the inclusion of a black sector in order to secure a satisfactory match. We may therefore make the deduction that, in the diagram of Figure 45, the place for the color which bears the name, unqualified, of yellow is on the line drawn from W through the mid-point of RG, but two or three times as distant from the latter point. The extraordinary luminosity of yellow, unequalled by any other color of as high a degree of saturation, is explained by a prismatic analysis of the color, which shows that such yellow is merely white deprived of its blue and violet components, both of which are of relatively feeble luminosity. This experiment also proves the converse, that the colors last named are relatively strong colors, that is, that a small admixture of either of them would produce an inordinately great change in the hue. In accordance with the latter deduction the spectrum of a deep blue sky exhibits astonishingly little difference from that of a white cloud.

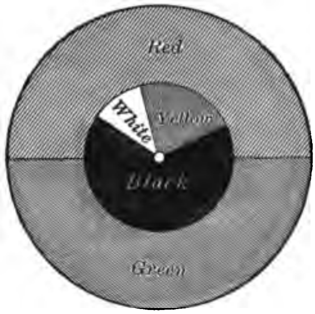

FIGURE 46.

Although a large number of names has appeared in the

discussion of the color diagram, there are certain groups of names which have not appeared at all; such are the browns, olives, lavenders, etc. For a physical definition of these terms we must look to variation in the elements of luminosity and of saturation. A simple experiment will show that any hue lying between red and yellow, of a high degree of saturation and a very low degree of luminosity, is a brown. Hues from greenish yellow to yellowish green with the same characteristics yield the group of olives. These statements may be readily tested by putting a small sector of bright-colored paper of the hue in question upon a disk of black paper and rotating it rapidly. By a similar experiment with any hue from pure green to magenta and even somewhat beyond, we meet with no new color terms, but content ourselves with the prefix "dark" as sufficiently descriptive. Now vary the experiment by placing the narrow sectors of colored papers over white disks and set them in rotation; we shall then find colors which would be characterized as pale red, pale yellow, pale green, etc., until, having passed the region of pale blue, we find that ultramarine-blue thus diluted yields a lavender, while further progress in the same direction gives lilacs and pinks. It is an odd fact that in the first region described new color terms appear only when the luminosity is diminished and not when the saturation is decreased, while in the second region the exact contrary holds true. There is undoubtedly a reason for this, founded upon the psychology of color sensations, but it would probably be difficult to formulate it.

The natural colors of objects are due to their property of reflecting only a portion of the light waves falling upon them, and a portion differing greatly with different wavelengths. There are a few substances in which this action is confined to the surface, such as colored metals and a variety of less well-known things which we describe as possessing a metallic color. Certain aniline dyes in the solid state are good examples of the latter group of bodies. But in a vast majority of cases the modification of the reflected light is

produced by absorption within the material of the body. The light reflected from the surface of a ruby or of a sapphire is as white as that reflected from a diamond, but such light as has passed through any portion of the former gems appears of the characteristic color on account of the suppression of certain portions of the white light by absorption. Most pigments are simply small fragments of highly colored crystals imbedded in a transparent medium. Chrome yellow, for example, when used as a paint, consists of finely powdered crystals of lead chromate mixed with oil. If, by means of a prism, we examine light which has been modified by passing through a sheet of yellow glass, we shall find that the modification consists in a relatively great absorption of the blue and violet lights, and that it transmits the remaining colors of the spectrum quite completely. A similar test of a blue shows that it very strongly absorbs red, orange, and yellow, but is quite transparent to the remaining colors. If one of these colored glasses is placed over the other, and ordinary white light is directed through both, the transmitted light will be green, this being the only color to which both are transparent. This is very like the experiment which the painter makes when he mixes a yellow with a blue pigment and spreads the mixture upon a surface. From such a painted surface, light would come to the eye which may be traced to four sources: First, that which is reflected directly from the outer surfaces of the pigment grains; second, that which has been modified by absorption through yellow particles alone; third, that which has been affected by the blue particles and not the others; and, finally, that which has suffered successive absorption by both yellow and blue grains. Of these four varieties of light the first is white, the second and third, being complementary, combine to form white, and the fourth is left to determine the hue, which, as has been shown, is green.

When the intensity of a colored light is greatly increased, the resulting sensation always diminishes in saturation. This fact is found from observation, and is hardly an immediate

consequence from the theory, although a plausible explanation might be deduced from it. On the other hand, should the brightness of a colored light be continuously decreased, it is found that the luminosity falls off at a rate differing greatly with different colors. Red, for example, loses its brightness with faint illumination very rapidly, whereas blue lies at the opposite extreme in this particular. In this physiological fact is found the explanation of numerous familiar phenomena. We may cite the prevalent bluish tone of a moonlit landscape as an example; also the surprisingly dark brown color of an orange in fading daylight.

It has been already stated that, as far as purely physical considerations are concerned, any three colors which may be so combined as to yield white might be made the basis of a color diagram. The particular set chosen by Young has the advantage of giving sensation curves which have but a single maximum (see Figure 44), but the decisive reason for choosing them is found in the peculiarities of vision exhibited by color-blind persons. If, for example, we suppose either one or two of the three fundamental color-sense organs to be defective or wholly wanting, a perfect theory ought to suffice to explain the differences between the color sensations of such a defective eye and those of the normal eye. The general success attending the application of the theory as stated, which would be quite lacking in another essentially different assumption for the fundamental colors, is a strong support in favor of the approximate truth of the theory.

We deduce from the theory that an eye incapable of recognizing a red sensation ought to find matches for all its color sensations in a proper admixture of green and blue, diluted with white or darkened if necessary. So, too, an eye wholly insensitive to green ought to be satisfied with a linear color diagram containing red and blue only. To the former of these defective eyes blue-green, being complementary to red, ought to appear indistinguishable from white, or at least from the white of low luminosity which we call gray. In exactly the same way we infer that the other defect would

carry with it the inability to distinguish a difference between magenta (the complementary color of green) and gray. These are only special cases among a large number of predictable phenomena from the theory, but they will serve our immediate purpose. Such defective color sense is far from uncommon, at least among men of whom we may perhaps reckon four in every hundred as more or less red-blind; green-blindness is far less frequent, and blue-blindness, although not wholly unknown, is of extreme rarity. Color-blindness is relatively extremely uncommon among women.

It is often desirable to determine the character of color vision in persons who are quite uneducated and who do not know the names of more than three or four colors. Holmgren has devised a remarkably simple and effective method of making such determinations, wholly independent of the power possessed by the individual to name colors correctly, and founded upon the theory which has occupied our attention. It may be easily described and easily applied. The examiner provides himself with a large number of skeins of worsted representing a wide range of colors, not only as regards hues, but also in respect to saturation and luminosity. One of these skeins is of a magenta color, and it is at first kept hidden from the observer. The person to be tested is shown a skein of green worsted, and is requested to select from the purposely confused collection all those which have the same color, irrespective of likeness in regard to saturation and luminosity. If the task is performed without error, the evidence of normal sense is almost absolute; but if grays or browns or other colors are selected as belonging to the same class as the green, abnormal color sense, or color-blindness, is demonstrated. It remains to find the character of the defect. In the first place, one may assume that it cannot be blue-blindness on account of its extreme rarity; it is therefore either red- or green-blindness. The next step in the test is to exhibit the skein of magenta worsted, and request that all skeins of its color should be chosen. If the eye examined is blind to red, only the blue present in the

magenta produces a color sensation; consequently violets and blues will be selected as allied in color to the magenta. If, on the other hand, the color-blind person finds in grays a match for the magenta, he betrays the possession of green-blindness, since the magenta is complementary to green.

When the consequences of the theory are pursued to the extreme, however, especially in the case of the faintest visual perceptions, it proves less satisfactory. For instance, if in a perfectly dark room one observes a piece of iron under continuously rising temperature, the first visual impressions ought to be, according to the theory, those of redness, since the first copious radiations are of long wavelength; but experience shows that the iron becomes visible long before any trace of color can be detected. Another fact bearing upon this point is that even the normal eye is color-blind for all objects very remote from the centre of most distinct vision. The only reason for this singular fact being wholly overlooked is because few people pay any careful attention to the character of oblique vision. Light sensations do not, therefore, necessarily involve a sensation of color. A peculiarity of the construction of the retina has suggested an addition to the theory. It is found that the cones are most abundant in the macula fovea, or region of most distinct vision, even to the exclusion of the rods altogether, while away from this locality the rods become more and more abundant up to the limit of the visual field, where the cones are very infrequent. If we assume that the seat of color sensation is in the cones, while a stimulation of the rods gives rise to a sense of brightness only, we shall have a theory somewhat better adapted to explain known phenomena, and one which has considerable support among physiologists.

CHAPTER IX

THEORIES CONCERNING THE NATURE OF LIGHT

According to the views of Newton, visual sensations were produced by minute corpuscles projected with enormous velocities from luminous bodies, differences of color being due to differing size in these minute bodies. When these corpuscles approach the boundary of an optically denser medium, they are subjected to a force of attraction which causes them to deviate from their otherwise rectilinear paths. This is the explanation of the phenomenon of refraction. The secondary phenomenon of dispersion was simply and naturally explained by an assumption that this attracting force varies with differing size. Singularly enough, the explanation of one of the most common phenomena, that of partial reflection at the boundary of a transparent medium, offered formidable difficulties: How is an attraction which is necessary to account for refraction also to act as a repulsive force in the case of a portion of the corpuscles? This is a difficulty which the advocates of Newton's theory have never been able to meet in a satisfactory manner.

When Newton attempted to extend his theory to the explanation of the colors of thin plates, a subject which he was the first to investigate in a scientific manner, it was found even less satisfactory. He was obliged to supplement it with the highly artificial hypothesis that the corpuscles experience periodic changes in the ease with which they enter a refracting body. Even this addition to the theory fails to yield more than a rough approximation to an explanation of the phenomena, since the blackness of the central spot in Newton's rings apparatus, when the plates are brought

into contact, is in contradiction with it. But it was only on account of a subsequently accumulated knowledge of optical phenomena which refused to adjust themselves to this theory, no matter how it might be modified, that its final overthrow came about. This not only required a long time, but also a champion of transcendent power to break with a system which had the force of tradition as well as the authority of the greatest of all philosophers to support it.

From 1704, the date of the publication of Newton's Optics, until 1815, the corpuscular theory was hardly questioned; at any rate, it reigned supreme in all treatises on light, and was questioned only by a very few investigators, who failed to acquire an influence that was anywhere decisive. In the latter year a remarkable man, Augustin Fresnel, a young French engineer in government employ, entered upon a career of scientific activity which proved of almost unprecedented brilliancy and success. This, as far as it bears upon the purely physical theory of light, may be regarded as completed in 1826. Beginning with a highly philosophical criticism of some of the teachings of the accepted doctrines of optics, supported by the most apt appeals to experiment, he quickly extended his investigations until they embraced nearly all phenomena of light known to his contemporaries; and with such success that he established as beyond question the essential truth of a wave theory, bringing it to so high a degree of completion that his views were long supposed to be practically final. On account of the importance of this work of Fresnel in the history of physical science of the past century, it is well worth our while briefly to review his achievements.

The phenomena of diffraction first engaged the attention of Fresnel. It had long been known that the shadow of an opaque body cast by a point-source of light is somewhat different from what would be supposed from simple geometrical considerations, the difference consisting chiefly in an encroachment of the light upon the borders of the shadow. Newton, who called this phenomenon inflection, attributed it

to an attractive force exerted by the opaque body upon the corpuscles while in its neighborhood, thus causing an in-bending of their paths. Fresnel showed that this theory was quite untenable, since the inflection caused by the back of a razor is exactly the same as that caused by the edge, although in the former case it is manifest that the time during which the corpuscles are subject to the deflecting force is far greater than in the latter. By similar appeals to ingenious crucial experiments he demonstrated that none of the current theories was sound; but far from resting here, he showed that all the observed phenomena could be perfectly accounted for in the undulatory theory of light, by an application of the principle of Huyghens. Extending this principle, so fertile in his hands, to wider fields in the domain of optics, he found in every case that the new method was adequate to yielding perfectly satisfactory explanations. Starting with quite simple and natural hypotheses as to the conditions which must exist at the common boundary of two transparent media, he was even able to deduce quantitative laws governing the phenomena of reflection and refraction, which accord surprisingly well with experiments devised to test them.

A few years before the commencement of Fresnel's activities, Malus, while looking through a double-image prism, observed that the two images of a distant window, which happened to be in such a position as to reflect light strongly to his eye, were quite different in brightness, and under some circumstances one image might entirely disappear. Further study showed that all ordinary transparent substances were capable of thus modifying light by reflection, and that the modification is complete at an angle which is simply related to the refractive index; moreover, that under the latter conditions the light would be transmitted through a doubly refracting crystal in certain directions without bifurcation. Such modified light is called polarized light, and the phenomena described are two of the simplest of an enormously extensive class, many of which are of extraordinary beauty. This discovery and those which quickly followed in the

same field presented a host of new and difficult problems to physicists. Of the many active and able workers in this domain Fresnel was easily the leader. In a very few years he had proposed and developed a general theory of light which embraced these new phenomena and which stood almost unquestioned until our own day. Although this subject is far too extensive and intricate for any adequate presentation here, we may consider its most general outlines. Fresnel's theory supposes that the motion of the particles which constitute the vibrations of light is always in a direction at right angles to the line of propagation of the waves. When the paths of the particles are quite irregular or without order, the light is ordinary light; but when the paths are similar, whether straight lines, ellipses with their axes parallel, or circles with a common direction of motion, the light is said to be polarized. From this simple hypothesis he succeeded in erecting an extraordinary structure which harmonized and explained nearly every known phenomenon of light in a manner that until the most recent times practically withstood all destructive criticism. Even recent achievements in this domain of science have been supplementary to, rather than subversive of Fresnel's general work. Of the phenomena known to his contemporaries, that of dispersion alone was unconsidered by him, a phenomenon which obviously depends upon the ultimate molecular structure of the refracting substance and which has recently been reduced to comparatively simple laws. This great work of Fresnel was looked upon, as indeed it well deserves to be, as one of the greatest monuments to the human understanding — comparable to Newton's doctrine of universal gravitation — and it long remained of almost unquestioned authority. Ultimately, however, one of its fundamental postulates, namely, that the vibrations are always at right angles to the direction of the motion of the light, began to give rise to difficulties. The fact also that the theory could not determine specifically whether the direction of vibration of plane polarized light is in the plane of polarization or perpendicular to it was not

only a manifest incompleteness, but it was a constant stimulus to a critical inspection of its premises. The more these points were studied the more insoluble the difficulties appeared, until there came to be a tolerably wide-spread belief that the theory was not only incomplete, but that in some way it must be essentially in error. To acquire a notion of what modern science has done to clear up these points, we must first review a class of phenomena which seem to be totally unconnected with optics, but which in the end will afford a very remarkable example of the tendency of all science toward unity.

In 1845 Faraday discovered that if polarized light is passed through a transparent substance in a magnetic field and in the direction of the field, the plane of polarization is rotated. The amount of rotation for any given substance is found to be proportional to the strength of the magnetic field and to the length of the path in the material. As many substances, such as turpentine, a solution of common sugar, quartz crystals in the direction of their crystalline axes, etc., present us with a similar fact, this would not be so surprising save for a remarkable difference in the two cases which may be thus described: When the plane of polarization is rotated by passing through a sugar solution or a similar body, and the transmitted light is reflected back upon its course so as to retrace its path, it is found that the original angle of the polarization is perfectly restored by a precisely equal rotation in the opposite direction in the return; but a similar experiment upon the body giving the magnetic rotation shows a doubled change of the angle. This indicates that, although in the first case we must explain the rotation by the molecular constitution of the material, we are not permitted to suppose that the magnetic field has produced a similar molecular structure in the second case, since the rotation is constant in direction irrespective of the direction of the motion of the light. Of course, from the known nature of magnetism, this is equivalent to asserting that there must be some relation between light and electricity. But this is not the

most obvious connection between these two classes of phenomena, for, as we now know, the earliest division of materials in accordance with their electrical properties involved a classification according to their most characteristic optical properties also. Thus all conductors of electricity, excepting only those liquids which undergo a chemical decomposition when they transmit an electrical current, and therefore belong to an obviously different class, are extremely opaque to light; conversely, all substances which are good insulators are also transparent to light, at least to an extent which would make a sheet a few hundred-thousandths of an inch in thickness appear perfectly transparent, although such a sheet of any metal or similar conductor would not differ greatly in opacity from a thick plate. An excellent illustration of the generality of this law is furnished by the element carbon, which in the dense opaque form — like graphite, for example — is a very good conductor of electricity, but in the form of the transparent diamond is an insulator.

Before the middle of the past century two methods of measuring electrical magnitudes had been developed; one of these is based upon the repulsion which exists between two electrically charged bodies, and the other upon the repulsion which exists between two similar magnet poles. Elaborate and repeated investigations have demonstrated that if a given electrical magnitude is measured according to one of these systems, and the value thus found is compared to a measurement of the same quantity in the other system, the ratio involves a velocity only. This statement is quite independent of the kind of magnitude chosen for the experiment. Within the limit imposed by unavoidable errors of observations, the value of this velocity always appears to be the same as the velocity of light.

Here, therefore, are three distinct relations between light and electricity, which have been long known, to no one of which is it possible to attach any *a priori* reason. It was left to Maxwell to illuminate this obscure field. His long and successful investigations in electricity and

magnetism, especially his efforts to reduce Faraday's brilliant discoveries to correlation and to consistent scientific statement, led him to the conclusion that light itself consists of electrical vibrations. He attempted to test the validity of this hypothesis by every means at his command. For example, according to his theory a non-magnetic substance ought to have a dielectric constant, or what Faraday named its specific inductive capacity, proportional to the square of its index of refraction. This indicated relation was found to hold with all expected precision in some cases, but to be widely removed from the truth in others. Again, since, according to the theory, only those substances are transparent which will offer a resistance to the motion of electricity within them analogous to elastic reaction, there ought to be a determinable relation between electrical conductivity and opacity. Maxwell attempted to find this relation in the case of gold-leaf, which is sufficiently thin to transmit a measurable portion of light falling upon it. Notwithstanding that the discrepancy was here found disappointingly great, the gradual accumulation of knowledge of the more recondite phenomena of the electrical field had led the great majority of physicists to the conclusion that Maxwell's theory was at least a close approximation to the truth, and accordingly one of the most brilliant discoveries of the nineteenth century. This may be regarded as a fair statement of the attitude of the world of science in 1888, when Hertz, a German physicist of extraordinary merit, whose early death was a great loss to science, made a series of remarkable experiments which have eliminated all possible doubt as to the essential verity of Maxwell's theory of light. Fortunately it is not difficult for us to gain a sufficient knowledge of the character of these experiments to enable us to understand their general bearing.

It had long been known that a Leyden jar suddenly discharged through a thick wire gives rise to an oscillatory current of very brief duration, and that in certain simple cases the period of the oscillations can be calculated with considerable precision. Hertz recognized that during the

time of discharge such an electric circuit must be a source of oscillatory changes in the magnetic field, which, if the views of Maxwell are in accordance with fact, should be propagated through space with the velocity of light. Although it is difficult, if not quite impossible, to measure directly this velocity, if one knows the wavelength and the period it is perfectly easy to deduce the velocity from these two elements, since in its period every wave moves a distance equal to its own length. In these experiments the period was found by calculation from the elements of the electric circuit; it only remained therefore to determine the length of the waves. Hertz accomplished this in the following simple and ingenious manner: At a considerable distance from the source of the waves he placed a large sheet of metal perpendicular to its direction from the source. From this sheet the waves were returned upon themselves by reflection. Now, a well-known fact in wave motions is that when two systems of waves of like period are moving in opposite directions, they combine to form a system of standing waves of half the length of the free waves. The regions where motion is destroyed by this kind of interference are called nodes. Hertz demonstrated the existence and positions of these nodes by means of an apparatus which possessed the same electrical period as his source. This apparatus he called a resonator. The value of the velocity of these waves deduced from his observations differs no more from the known velocity of light than would be expected from the unavoidable errors of observation; thus it complies with the requirements of Maxwell's theory. These waves, therefore, are shown to differ from light waves only in their enormously greater wavelengths, and they must be subject to all the established laws of optics which are independent of the length of the waves. The last conclusion was thoroughly tested by Hertz by a series of most interesting and convincing experiments. He found that strictly according to the laws of optics these waves are reflected from the surfaces of all bodies which conduct electricity; that they readily pass through substances which

behave as insulators; and that in passing from one insulating medium to another the direction of propagation is altered in accordance with the law of sines. Further than this, he showed that such electrical waves admit of polarization, and they are therefore characterized by motions at right angles to the direction of propagation.

During the time which has elapsed since these investigations, a host of experimenters have improved the methods and apparatus of Hertz, and have largely extended the range of wavelengths that can be observed. At present these Hertzian waves are being applied to transmitting electrical signals to distant points without the use of connecting wires. On the other hand, many investigators have been employed in the application of analysis to both the old and the new problems in optics. The difficulties which attach to Fresnel's mode of regarding the optical phenomena of crystalline media are found to disappear, and all the complex phenomena of light admit of explanation from a consistent body of premises.

This great advance, however, cannot be regarded as an unmixed gain, since it replaces comparatively simple mechanical considerations, from which we can construct mental pictures of the phenomena in a medium submitted to the action of radiant energy closely allied to those which are suggested by ordinary experience, by conceptions of agencies that no philosopher has yet succeeded in grasping even to an extent necessary to form a clear notion of what constitutes an electrically charged body; much less what an electric vibration may be. It is not surprising that the earlier views are the simpler ones, else they hardly would have had precedence in time; but there is no possible doubt that no purely mechanical theory has been formulated, even after a century of effort, which is proof against all criticism, nor, on the other hand, has any one hitherto been able to urge an objection to the electrical theory now prevalent, save the difficulty of understanding its fundamental conceptions. We may hope to have some time, perhaps at no very distant

future, a popular treatise on light which shall start from the phenomena presented by a rubbed stick of sealing-wax, and by logical development deduce all the phenomena of light now known, and very likely a host of others not yet dreamed of.

APPENDIX A

The usual formulas for the calculations of lens systems other than those of the greatest simplicity are extremely complicated. Should such accuracy be desired as is absolutely requisite for the designing of optical apparatus, in which case the ordinary approximations are insufficient, the formulas become almost unmanageable.

The famous paper by Gauss, which was published in 1840, and which first introduced the conception of cardinal points in an optical system, was a very remarkable advance on the older theories. After its appearance it was possible to take into account the thicknesses and separations of the lenses with rigid accuracy, without introducing greater complexity in the final equations than were already secured by a method which assumed that these elements might be neglected. But even more important than this was his perfect definition of terms which had been used for ideal systems, but which could not have any exact meaning when applied to any physical apparatus, such as focal length, magnification, optical centre. This paper, with Listing's addition of the two nodal points, left nothing to be desired in giving a complete and definite geometrical notion of an optical system whose cardinal points were known. But this is not the primary end either in designing an optical system or in determining its optical efficiency when completed. The absolute focal length and magnification of any optical instrument is of the least possible consequence, as becomes at once obvious when we consider how few working astronomers or microscopists can give these data with precision for their own instruments.

What is necessary in designing an optical instrument is a knowledge of its aperture, in the more general sense, of the variation of its focal plane with varying wavelength of light, of its spherical aberration, of the variation of the spherical aberra-

tion with varying wavelength, and, if the character of the image remote from the axis is of consequence, the variation of the magnification with varying wavelength, and, finally, its astigmatism. Neither the older methods nor that of Gauss defines the first of these important quantities, except, of course, in the simple case of the telescope, nor are they convenient for calculating the others.

From the character of the quantities named above, it is obvious that their mathematical expressions must be for the most part of the nature of derivations from the fundamental equations. But all the fundamental equations for lenses of sufficient exactness to be of the least value are so complicated that their derivations, at least, are quite valueless for practical purposes.

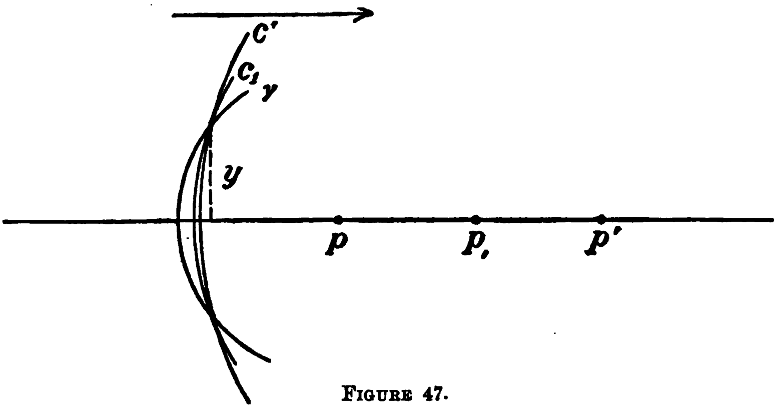

FIGURE 47.

The use of the natural conceptions of wave-surfaces, of refracting surfaces characterized by their curvatures, and of light velocities, which has been employed in the text of this book, may well possess a commensurate advantage when reduced to rigidly quantitative form. Indeed, I have myself been well content with this innovation during many years of use in practice. The following sections may be looked upon as a brief exposition of the efficiency of the method and as a proof of all the principles employed in the text. It should be added that nearly all of this appendix is a slightly modified reproduction of a paper in the Memoirs of the National Academy for 1893.

I. GENERAL EQUATIONS

Let p, p', and p_1 be, respectively, the geometrical centres of the lens surface, the incident wave-surface, and the refracted wave-surface; also let γ, c', and c_1 be the curvatures of these surfaces in order, a positive curvature corresponding to the case where the centre of curvature is in the direction of propagation of light from the surface. Let x, x', and x_1 be the sagittas of these surfaces in order named, corresponding to the common semi-chord y; then for small values of y we have

$$x = \tfrac{1}{2}\,y^2\gamma, \qquad x' = \tfrac{1}{2}\,y^2 c', \qquad x_1 = \tfrac{1}{2}\,y^2 c_1.$$

From the laws of wave propagation we have, if ρ represents the ratio of the velocity in the medium to the right of the surface γ to that in the medium to the left,

$$x - x_1 = \rho\,(x - x')\,.$$

Substituting the above values of x, x', x_1 in this equation, eliminating the common factor $\frac{1}{2}\,y_2$, and solving for c_1, we have

$$c_1 = \gamma\,(1 - \rho) + \rho c'.$$

This wave-surface will be propagated with uniform velocity and uniformly decreasing radius until it reaches the second refracting surface at a distance t from the first, when its radius is reduced from $\frac{1}{c_1}$ to $\frac{1}{c_1} - t$, and therefore its curvature to the reciprocal of this, or to

$$\frac{c_1}{1 - c_1 t} = \mu' c_1$$

if we define μ' by this equation. After refraction at this second surface, the curvature will be according to exactly the same reasoning as that applied to the first surface,

$$c_2 = \gamma'\,(1 - \rho') + \mu'\,\rho' c_1\,.$$

This completes the general solution, and it can be extended to any number of refractions, thus:—

$$\begin{array}{ll} c_1 = \gamma\,(1-\rho) + \mu\rho c' & \mu = 1 \\ c_2 = \gamma'\,(1-\rho') + \mu'\rho' c_1 & \mu' = (1 - c_1 t_1)^{-1} \\ c_3 = \gamma''\,(1-\rho'') + \mu''\,\rho'' c_2 & \mu'' = (1 - (c_2 t_2)^{-1} \quad (a) \\ \ldots\ldots\ldots & \ldots\ldots \\ c_\lambda = \gamma^{\lambda-1}\,(1-\rho^{\lambda-1}) + \mu^{\lambda-1}\,\rho^{\lambda-1}\,c_{\lambda-1} & \mu^{\lambda-1} = (1 - c_{\lambda-1}\,t_{\lambda-1})^{-1} \\ c_{\lambda+1} = \gamma^\lambda\,(1-\rho^\lambda) + \mu^\lambda\,\rho^\lambda\,c_\lambda & \mu^\lambda = (1 - c_\lambda\,t_\lambda)^{-1} \end{array}$$

If the lenses are all indefinitely thin and in contact — in other words, if all the t's are equal to zero, and therefore all the μ's equal to unity — the equations (a) can be readily combined into one. This reduction, by inspection, is

$$\begin{aligned} c_{\lambda+1} = \gamma^\lambda\,(1-\rho^\lambda) + \gamma^{\lambda-1}\,(1-\rho^{\lambda-1})\,\rho^\lambda + \ldots\ldots \\ + \gamma\,(1-\rho)\,(\rho^\lambda\,\rho^{\lambda-1}\ldots\ldots\rho') + \underline{\rho^\lambda}\,|\,c'. \end{aligned}$$

If n_0 is the index of refraction of the first medium, n' in that of the second, and so on, we shall have

$$\rho = \frac{n^\circ}{n'},\ \rho' = \frac{n'}{n''}\quad \ldots\ldots \quad \rho^\lambda = \frac{n\lambda}{n^{\lambda+1}},$$

and the product of all these quantities will be equal to $\frac{n_0}{n^{\lambda+1}}$, which quantity, the ratio of light velocity in the last medium to that in the first, we will designate by ρ_0. By this substitution the above equation becomes

$$c_{\lambda+1} = \mathrm{P} + \rho_0\,c',$$

where P is a constant, depending solely upon the physical constants of the system; it is called the *power* of the combination, not only because it is the change in curvature which the system can produce in a plane incident wave-surface ($c' = 0$), but also because in the most common of all cases in which the first and last media are alike, and therefore $\rho_0 = 1$, P is the curvature which the system adds to every incident wave-surface.

II. IMAGES AND MAGNIFICATION

Let o' (Figure 48) be the indefinitely small distance of the centre of the incident wave-surface c' from the axis, and o_1 that of the centre of the refracted wave-surface c_1; then we have, from the

general law of wave motion, the line o_1 is the image of o', and the angle which c_1 makes with γ at the vertex is ρ times that which c' makes at the same point; hence

$$\frac{o_1}{o'} = \rho\frac{c'}{c_1}.$$

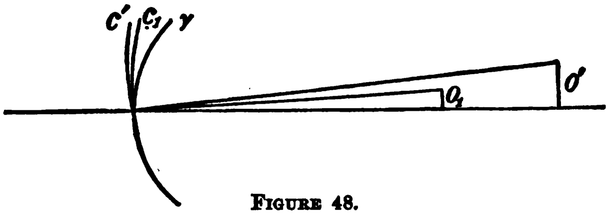

FIGURE 48.

The image of o_1 after refraction at a second surface we may call o_2; then by similar reasoning we have

$$\frac{o_2}{o_1} = \rho'\frac{\mu' c_1}{c_2}.$$

Extending this process to $\lambda + 1$ refractions and collecting, we have

$$\frac{o_1}{o_2} = \rho\frac{c'}{c_6}$$

$$\frac{o_2}{o_1} = \rho'\frac{\mu' c_1}{c_2}$$

$$\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot$$

$$\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot$$

$$\frac{o_\lambda}{o_{\lambda-1}} = \rho^{\lambda-1}\frac{\mu^{\lambda-1} c_{\lambda-1}}{c_\lambda}$$

$$\frac{o_{\lambda+1}}{o^\lambda} = \rho^\lambda\frac{\mu^\lambda c_\lambda}{c_{\lambda+1}}$$

The product of these equations is, remembering that the product of all the ρ's is ρ_0,

$$\frac{o_{\lambda+1}}{o'} = \rho_0\frac{c'}{c_{\lambda+1}}\ \underline{\mu^\lambda|}\,.$$

The value of the factor $\underline{\mu^\lambda|}$ is easily changed by substituting successively the value of μ^λ and of c_λ given in (a), to the form

$$\left.\underline{\mu^{\lambda}}\right| = \left.\underline{\mu^{\lambda-1}}\right| \frac{1}{1 - c_\lambda t_\lambda} = \left.\underline{\mu^{\lambda-1}}\right| \frac{1}{a_{\lambda-1} + \beta_{\lambda-1}\, \mu^{\lambda-1} c_{\lambda-1}},$$

in which $a_{\lambda-1}$ and $\beta_{\lambda-1}$ are constants depending solely upon the physical constants of the system, and not at all upon the value of any of the wave curvatures. If we make another substitution of the values of $\mu^{\lambda-1}$ and of $c_{\lambda-1}$ in this last equation, we have

$$\left.\underline{\mu^{\lambda}}\right| = \left.\underline{\mu^{\lambda-2}}\right| \frac{1}{a_{\lambda-2} + \beta_{\lambda-2}\, \mu^{\lambda-2} c_{\lambda-2}}.$$

If we repeat this process of substitution λ times, we shall have finally,

$$\left.\underline{\mu^{\lambda}}\right| = \frac{1}{a + \beta c'},$$

in which a and β are determinable functions of the γ's, the t's, and the ρ's, that is, of the physical constants of the system, and the ratio of the ultimate image to the object, or the magnification becomes

$$\frac{o_{\lambda+1}}{o} = \rho_0 \frac{c'}{c_{\lambda+1}\,(a + \beta\, c')}. \qquad (b)$$

There is, however, another very important expression for the magnification which may be derived as follows: —

Suppose we have a diaphragm at the first surface with the small diameter $2a$. Let the semi-angular diameters of the wave-surfaces, incident and refracted, thus limited be designated by ω' and ω_1, and so on for the successive refracting surfaces; then —

$$\begin{aligned}
\sin \omega' &= ac' \\
\sin \omega_1 &= ac_1 \\
\sin \omega_1 &= a_1 \mu' c_1 \\
\sin \omega_2 &= a_1 c_2 \\
\sin \omega_2 &= a_2\, \mu''\, c_2 \\
\sin \omega_3 &= a_2\, c_3 \\
&\cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\
\sin \omega_\lambda &= a^\lambda \mu^\lambda c_\lambda \\
\sin \omega_{\lambda+1} &= a_\lambda\, c_{\lambda+1}
\end{aligned}$$

From these we have

$$\frac{\sin\omega'}{\sin\omega_1}=\frac{c'}{c_1} \qquad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot$$

$$\frac{\sin\omega_1}{\sin\omega_2}=\mu'\frac{c_1}{c_2} \qquad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot$$

$$\frac{\sin\omega_2}{\sin\omega_1}=\mu''\frac{c_2}{c_3} \qquad \frac{\sin\omega^{\lambda}}{\sin\omega_{\lambda+1}}=\mu^{\lambda}\frac{c_{\lambda}}{c_{\lambda+1}}$$

The product of these equations gives

$$\frac{\sin\omega'}{\sin\omega_{\lambda+1}}=\frac{c'}{c_{\lambda+1}}\,\mu^{\lambda}\Big|$$

Multiplying this equation by ρ_0 and observing (b) we have

$$\frac{o_{\lambda+1}}{o'}=\rho_0\frac{\sin\omega'}{\sin\omega_{\lambda+1}} \qquad (c)$$

This equation has been proved for small values of a only, and would therefore hold good if either ω or $tg\ \omega$ were substituted for $\sin\omega$. In the last form, $tg\ \omega$ replacing $\sin\omega$, the equation was first given by Lagrange, and was a most important contribution to the theory of optics; but it is possible to show that the form

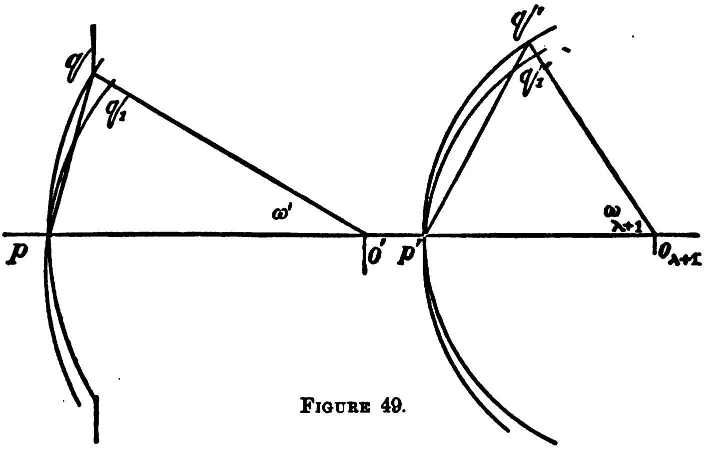

FIGURE 49.

which we have given above is rigidly true, independently of the size of a, provided only that the incident and finally refracted wave-surfaces are both spherical, that is, that the system is free from spherical aberration.

Let pq (Figure 49) be the incident wave c', and $p'q'$ the finally refracted wave-surface $c_{\lambda+1}$. If c' is limited by a diaphragm at q, then $c_{\lambda+1}$ is also limited at a point q', which is the *corresponding* point to q; that is to say, the point in $c_{\lambda+1}$ where all the light energy comes from a region in c' indefinitely near q. Call the semi-angular aperture of c' and of $c_{\lambda+1}$, ω' and $\omega_{\lambda+1}$, as indicated in the figure. Now suppose the incident wave to be inclined by an indefinitely small angle $o'c'$, then the finally refracted wave will also be inclined by an infinitely small angle at p', which will be equal to $o_{\lambda+1}\, c_{\lambda+1}$. It is apparent from the figure that in the new oblique waves q_1 is a *corresponding* point to q, since the wave-surface is propagated in the direction of its normal, and, for the same reason, q_1' is a corresponding point to q', hence q_1' on the finally refracted oblique wave is a corresponding point to q_1. But since p and p' are corresponding points for both wave systems, the time required for light waves to pass from q to q' and from q_1 to q_1' is, in each case, equal to the time required in going from p to p'; hence the time required for progression from q to q_1 is equal to the time from q' to q_1'. The velocity of propagation in the last medium, however, is ρ_0 times as great as in the first; consequently we have

$$\rho_0 \cdot qq_1 = q'\, q_1'.$$

By inspection of the figure we see that

$$qq_1 = pq \cdot o'c' \cos \tfrac{1}{2}\, \omega'$$
$$q'q'_1 = p'q' \cdot o_{\lambda+1}\, c_{\lambda+1} \cos \tfrac{1}{2}\, \omega_{\lambda+1}\,.$$

Combining these three equations and substituting the trigonometrical expression for the chords pq and $p'q'$, we have

$$2\rho_0\, o' \sin \tfrac{1}{2}\, \omega' \cos \tfrac{1}{2}\, \omega' = 2\, o_{\lambda+1} \sin \tfrac{1}{2}\, \omega_{\lambda+1} \cos \tfrac{1}{2}\, \omega_{\lambda+1}\,,$$

whence we derive immediately the equation (c).

This highly important relation is essential in calculating the absolute optical power of every optical apparatus except the telescope. Its truth was assumed by Professor Abbe in his celebrated paper on the limit of power in microscopes, and was proved very indirectly by Helmholtz in his paper on the same subject as a consequence of the second law of thermodynamics. I am not aware that any proof purely optical has heretofore been given.

III. SIMPLIFICATION OF GENERAL EQUATIONS (a)

In the discussion to this point we have made c' the curvature of the incident waves at the first vertex of the system, and $c_{\lambda+1}$ the curvature of the finally refracted waves at the last vertex. We will now seek the relation of the curvatures for incident, and finally refracted wave-surfaces at points situated, respectively, at a distance x_0 from the first vertex and x_1 from the last.

Let us suppose that the incident wave-surfaces are bounded by a circular diaphragm of radius a_0 at x_0, then the finally refracted wave-surfaces will be also bounded and will have a definite semi-diameter at x_1, which we shall designate by $a^0{}_1$; then, if C' and C_1 represent the wave curvatures at the points x_0 and x_1 for incident and finally refracted waves, respectively, we have the following relations: —

$$\sin \omega' = a^0 C' \qquad\qquad c' = \frac{C'}{1 + C'x_0}$$

$$\sin \omega_{\lambda+1} = a^0{}_1 C_1 \qquad\qquad c_{\lambda+1} = \frac{C_1}{1 + Cx_{11}}$$

Substituting these values in (c), we have

$$\frac{o_{\lambda+1}}{o'} = \rho_0 \frac{a^0 C'}{a^0{}_1 C_1}$$

By replacing the left member of this equation by its value given in (b) and then substituting the values above for c' and $c_{\lambda+1}$ we have

$$\frac{a^0}{a^0{}_1} = \frac{1 + C_1 x_1}{\alpha + (\alpha x_0 + \beta)\, C'},$$

whence the value of C_1 is given by the equation

$$C_1 = \frac{\frac{a^0}{a^0{}_1}\alpha - 1}{x_1} + \frac{\alpha x^0 + \beta}{x_1}\frac{a^0{}_1}{a^0} C'$$

If x_0 and x_1 are so chosen that one is the optical image of the other, $\frac{a^0}{a^0{}_1}$ becomes a constant since $a^0{}_1$ is simply the image of a^0, and the expression for C_1 becomes greatly simplified. We

will call the value of $\frac{a^0}{a^0_1}$ for any such case k. This value can be found by substituting $\frac{1}{x^0}$ and $\frac{1}{x_1}$ for c' and $c_{\lambda+1}$ in equation (b), whence, remembering that a^0 corresponds to o', we have

$$\frac{a^0_1}{a^0} = \frac{1}{k} = \rho_0 \frac{x_1}{ax^0 + \beta}.$$

When $C'' = 0$ the value of C_1 is $\frac{ka - 1}{x_1}$; but with this value of C' we have likewise $c' = 0$, since both equations mean that the incident wave-surface is flat. Compute the value of $c_{\lambda+1}$ for this value of c' by means of equations (a) and call it $c^0_{\lambda+1}$; then from the equation above connecting $c_{\lambda+1}$ and C_1, namely,

$$c_{\lambda+1} = \frac{C_1}{1 + C_1 x_1},$$

we find
$$\frac{ka - 1}{x_1} = kac^0_{\lambda+1}.$$

Substituting these expressions in the general equation, we have this remarkably simple expression to replace equations (a),

$$C_1 = kac^0_{\lambda+1} + \rho_0 k^2 C''. \qquad (d)$$

Before discussing the methods for finding the contents, k, a, and $c^0_{\lambda+1}$ in this general equation, we will first establish expressions for the magnification and for the change in direction of the wave-surface in progressing from x_0 to x_1.

The term magnification has two distinct meanings, namely, the ratio of the dimension of the images to that of the object measured at right angles to the axis, and a corresponding ratio in the direction of the axis. The first we may call the transverse magnification and designate M; the other may be called the longitudinal magnification and designated L.

From the fifth equation of this section we may write at once the value for the first species of magnification. It is equal to

$$M = \rho_0 k \frac{C''}{C_1}. \qquad (e)$$

The longitudinal magnification is obviously equal to the ratio of the displacement of the image along the axis to the corresponding displacement of the object. From this definition and equation (d) we find

$$L = \frac{d\,\frac{1}{C_1}}{d\,\frac{1}{C'}} = \frac{d\,C_1}{d\,C'}\,\frac{C'^2}{C_1^2} = \rho_0 k^2\,\frac{C'^2}{C_1^2}.$$

From these two equations we find

$$L = \frac{M^2}{\rho_0}. \tag{f}$$

This law explains why the depth of field in microscopic vision seems so small, and also why an object under the microscope appears so much flatter when mounted in a medium of high refractive power, for in this case ρ_0 is greater than unity.

A consideration which is of much importance is the relation of the directions of the wave-surfaces. This relation is readily determined for the points x_0 and x_1. In Figure 49 let p and p' be the points x_0 and x_1, respectively; then, since by definition p' corresponds to p, we need make no restrictions as to the value of the angles of inclination of the oblique wave-surfaces. Call these angles ϕ_0 and ϕ_1 respectively; then, if pq is small, we have, as before,

$$\rho_0\, qq_1 = q'q'_1$$

and

$$qq_1 = pq \sin\phi_0$$
$$q'q'_1 = p'q' \sin\phi_1\ ;$$

and finally, since $p'q'$ is the image of pq,

$$\frac{pq}{p'q'} = k.$$

From these equations we read at once

$$\frac{\sin\phi_1}{\sin\phi_0} = k\rho_0. \tag{g}$$

IV. TO FIND THE VALUES OF THE CONSTANTS IN EQUATION (d)

There are two cases presented in practice: First, when all the constants of the optical system are given, and, second, when we can only depend upon measurements as applied to the system as a whole. We shall consider these two cases in turn.

Case 1. — All the constants of the system being known, compute $c_{0\lambda+1}$ by making c' equal to zero in equations (a); the value of a is the reciprocal of $\underline{\mu_\lambda|}$ in this computation.

To find k we either assume the value of x_0 and thence compute x_1 and k, or, assuming the value of k, compute x_0 and x_1.

For the first method make $c' = \frac{1}{x_0}$ in (a); the resulting value of $c_{\lambda+1}$ equals $\frac{1}{x_1}$, and, substituting these values of c' and $c_{\lambda+1}$ in (b), remembering that $(a + \beta c')$ is the reciprocal of $\underline{\mu_\lambda|}$, we find the reciprocal of k at once. The solution is therefore complete.

If k is assumed, we compute $c^0{}_{\lambda+1}$ and a as before, then

$$x_1 = \frac{ka - 1}{kac^0{}_{\lambda+1}}$$

whence we find x_0 from (a) by making $c_{\lambda+1} = \frac{1}{x_1}$ when $x_0 = \frac{1}{c'}$.

Case 2. — Should we desire to find the constants of (d) by experiment, proceed as follows: —

Choose any convenient point in the axis of the system for x_0, placing there an object, for example, a scale of equal parts; the image of this object will be at x_1 and the ratio of the size of the object to the image will be k. Then find the place on the axis of the image of an object at an indefinitely great distance in front of the system, that is, when $C' = 0$. The reciprocal of the distance of this point from x_1 equals $kac_{\lambda+1}$, and the problem is solved.

V. ON PARTICULAR VALUES OF k IN EQUATION (d)

Inspection of the equation (d) suggests at least four values of k, which make the equation of special simplicity. These values

are $1, -1, \frac{1}{\rho_0}$, and $\frac{1}{-\rho_0}$. Substituting these in turn, and replacing the diacritical marks of that equation by convenient symbols, we have

$$C_p = ac^0{}_{\lambda+1} + \rho_0 C^p \qquad (d^{\text{i}})$$
$$C_{-p} = -ac^0{}_{\lambda+1} + \rho_0 C^{-p} \qquad (d^{\text{ii}})$$
$$\rho_0 C_n = ac^0{}_{\lambda+1} + C^n \qquad (d^{\text{iii}})$$
$$\rho_0 C_{-n} = -ac^0{}_{\lambda+1} + C^{-n} \qquad (d^{\text{iv}})$$

The points defined by x_0 and x_1 when $k=1$ are called the first and second principal points, respectively. They were first introduced and their properties defined by Gauss, in a celebrated paper published in 1840. We see at once that the image of a small object at x_0 is at x_1, the transverse magnification is 1, and the longitudinal magnification is, from (f), $\frac{1}{\rho_0}$. The inclination of the incident wave at x_0 and the finally refracted wave at x_1 is given by (g), which becomes

$$\frac{\sin\phi_1}{\sin\phi_0} = \rho_0 .$$

For $k=-1$, the second of the above equations, the points x_0 and x_1 are called the first and second negative principal points. An object at x_0 has its image at x_1, the image being inverted, but of the same transverse dimensions as the object; the longitudinal magnification is the same as before, namely, $\frac{1}{\rho_0}$. Finally (g) gives

$$\frac{\sin\phi_1}{\sin\phi_0} = -\rho .$$

For $k=\frac{1}{\rho_0}$, the two points x_0 and x_1 are called the nodal points. These were first investigated and named by Listing in 1851. An object at the first nodal point has its image at the second nodal point, both axial and transverse magnifications being equal to ρ_0. These are the only two points so related that the image of a body at one point has the same shape and orientation as the body itself. A more important property is derived from equation (g), in which we see for this case

$$\frac{\sin\phi_1}{\sin\phi_0} = 1 ;$$

that is, a wave-surface which would pass through the first nodal point at an inclination ϕ_0 passes through the second nodal point under the same inclination after final refraction.

The final form, in which the points x_0 and x_1 may be called the negative nodal points, has the same longitudinal magnification for these points, but the transverse magnification and the relation of the inclinations are equal to those of (d^{iii}) taken negatively.

It will be observed that (d^{i}) is of exactly the same form as the equations for a system of infinitely thin lenses in contact. Moreover, if the first and last media are alike — in which case $\rho_0 = 1$ — equations (d^{i}) and (d^{iii}) become identical, as do also (d^{ii}) and (d^{iv}), or, in words, the principal and nodal points fall together, and also the negative principal and negative nodal points.

The determination of the position of all these points is the problem, when the constants of the system are known, of *Case 1* of the preceding section, and therefore need not be further discussed. But to determine them experimentally is not the same as *Case 2*, because they assume determinate values for k. We may proceed as follows: —

If both principal points are outside of the system and on opposite sides, we may find them at once by seeking the places of object and image when the image is erect and equal in size to the object. This process is applicable to some forms of compound microscope — such as that which is used as a terrestrial ocular in the telescope — but is very exceptional. In practically every other converging system the negative principal points will be outside the system and on opposite sides. Find these and the two principal focal points; then, since flat wave-surfaces have the same curvature at the second principal point as they do at the second negative principal point except with an opposite algebraic sign, as appears at once by making C^p and C^{-p} equal to zero in (d^{i}) and (d^{ii}), it follows that a principal focal point is exactly halfway between the principal point and the corresponding negative principal point.

The method of finding the four nodal points is precisely similar, except that the magnifications are taken as ρ_0 and $-\rho_0$ instead of 1 and -1, as in the case of the four principal points.

VI. MAGNIFICATION OF OPTICAL SYSTEMS

The general expression for the magnification is equation (*e*), which, of course, can be modified so as to be given in terms of the distance of the object from x_0, or from the first vertex of the system, or indeed from any other point fixed with respect to the system. But it is not easy to see that this would be of any practical interest. A photographer may desire a definite ratio of the image in his camera to the size of the object, and no doubt he could tell at once how far the object must be from the camera to give this ratio, if he had determined the value of k for two determinate positions of x_0 and x_1; but in practice his method of moving the instrument with respect to the object until the image becomes of the desired size would involve no more measurements than that of a single distance, which would be also necessary in the more recondite method.

For instruments used as aids to vision, however, the expression for magnification becomes particularly interesting, or rather the expression for angular magnification, since we care nothing for the absolute size of any image in question.

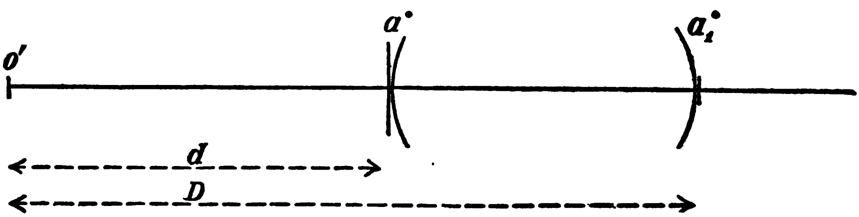

FIGURE 50.

Let Figure 50 represent any optical apparatus to be used as an aid in seeing o'. Let a^0 be an object in such a place that its image is at a^0_1 and very near the place of the eye. Call the distance from the eye to the object D, and the distance from a^0 to the object d, as represented in the figure. If d is very large, the instrument is called a telescope; if small, a microscope; when neither one nor the other, we have no name for it. For example, the instruments employed as optical aids in reading the distant circles of an equatorial may be called with equal propriety telescopes or

microscopes. In short, there is no precise distinction between the two types, and a general equation of the magnifying power ought to be applicable as well to one as the other.

We will define the magnifying power of such an instrument as the ratio of the apparent dimensions of a small object in the axis, as seen through the instrument, to that of the same object seen without the instrument.

From the diagram we see that the angular dimension of o', as measured from a^0 is $\frac{o'}{d}$, which is also twice the angle ϕ_0 for the point of the object furthest from the axis. The angular subtense of the object, as seen from a^0_1 through the instrument is $2\phi_1$, while the value of this angle without the instrument is $\frac{o'}{D}$, as is evident from the figure. This angle we call 2Φ. The magnifying power is therefore equal to

$$\frac{\phi_1}{\Phi}.$$

But from (g) we see that $\phi_1 = \rho_0 \frac{a^0}{a^0_1}\frac{o'}{d}$; whence, dividing by the value of Φ, the magnifying power becomes

$$\rho_0 \frac{a^0}{a^0_1}\frac{D}{d}$$

as a perfectly general expression.

For the telescope proper, $\rho_0 = 1$ and $\frac{D}{d} = 1$ also, since the length of the instrument is negligible with respect to the distance of the object. In this case the expression reduces to $\frac{a^0}{a^0_1}$, or, if we make a^0 equal to the diameter of the objective, it is equal to the quotient derived by dividing the diameter of the objective by the diameter of its image produced by the ocular, an old rule first given, I believe, by Ramsden.[1]

For the microscope the magnifying power is defined somewhat differently from our foregoing definition, namely, as the ratio of the apparent size to the size as seen at a distance of ten inches, which is $\frac{10^{\text{in.}}}{D}$ times as great. Thus modified, the expression becomes

[1] This will be recognized as the mathematical proof of the principle developed on page 115.

$$n \frac{a^0}{a^0_1} \frac{10^{\text{in.}}}{d},$$

since ρ^0 becomes the same as the index of refraction of the "immersion fluid" employed. Not only this, but we have also $\frac{1}{2}\frac{a^0}{d}$ equal to the tangent of one-half the so-called "angular aperture," whence the true aperture equals

$$n \sin \left\{ tg^{-1} \frac{a^0}{2d} \right\}.$$

It has long been known that the complete optical power of a telescope — that is, both its magnifying power and resolving power — could be determined by two linear measurements, the a^0 and a^0_1 of our discussion, but it has, perhaps, not been suspected before that with three linear measurements it is possible to determine both the magnifying power and resolving power of a microscope.

To illustrate the application of these formulas, we may quote the determinations of the optical constants of a Zeiss microscope employed with ocular No. 2, shortest tube, and three different objectives. The objectives were, a Zeiss *A*, of which his catalogue gives magnification 52 and numerical aperture 0.20, a $\frac{1}{4}^{\text{in.}}$ objective made by me, and a $\frac{1}{10}^{\text{in.}}$ Wales water-immersion. The measures were made by placing a glass scale upon the table of the microscope and bringing the objective into contact with it; the number of divisions of the scale visible above the ocular, and also the absolute length of its image were then recorded, the former length being a^0 and the latter a^0_1. Then the tube was raised a measured distance (d) until the scale was in focus. In the table the measures are in millimetres. *M* is the calculated magnification, and *A* the aperture.

Objective.	n.	a^0.	a^0_1.	d.	M.	A.
Zeiss *A*	1.00	3.5	2.0	8.5	52	.201
$\frac{1}{4}^{\text{in.}}$	1.00	3.3	1.7	3.5	140	.435
$\frac{1}{10}^{\text{in.}}$ immersion	1.33	1.1	1.5	0.48	351	0.99

It should be noted that, in order to determine *M* with precision, the ratio of a^0 to a^0_1, near the axis, should be taken, since the

surface a^0 of Figure 50 is not plane, but a portion of a wave-surface. For small apertures this distinction is insignificant.

To investigate the resolving power of an optical apparatus used in conjunction with the eye, we may proceed as follows: —

Suppose that we have a plane area at a distance of ten inches from the eye divided into regular spaces, and that we observe it through a hole of diameter a^0_1. The smallest angular values of the elements of which the field is made and which can still be distinguished as separate elements is about $4''.5 \frac{1^{\text{in.}}}{a^0_1}$, if a^0_1 is measured in inches. This is proved in all works on the wave theory of light and is in accordance with experience. Now this supposition corresponds precisely with the conditions of vision through an optically perfect apparatus of the type under consideration, save that the field is magnified in the ratio $\frac{a^0}{a^0_1}$ for the telescope and $n \frac{a^0}{a^0_1} \frac{10^{\text{in.}}}{d}$ in those systems where d is not indefinitely great, consequently in such cases the fineness of division of the field may be increased in these ratios. Hence the defining power of a telescope may be expressed by

$$4''.5 \frac{1^{\text{in.}}}{a^0_1} \frac{a^0_1}{a^0} = 4''.5 \frac{1^{\text{in.}}}{a^0}$$

and in the other class by

$$4''.5 \frac{1^{\text{in.}}}{a^0_1}, \frac{a_1{}^0 d}{n\, a^0\, 10^{\text{in.}}} = 4''.5 \frac{d}{10\, n\, a^0}.$$

The former of these equations is a familiar one, and need not be further discussed, but the latter contains the whole theory of the defining power of the microscope, and is therefore worthy of a brief consideration.

From what appears in the discussion of the relation of a^0_1 to a^0, we see that the former is the image of the latter, and also that a^0, being a portion of the incident wave-surface, is not plane as is, very nearly, a^0_1. We see, moreover, that the greatest possible diameter of a^0 is $2d$, in which case the incident wave-surface would be hemispherical; whence the maximum possible resolving power of a microscope is

$$\frac{4''.5}{20\, n}.$$

To reduce this to linear value we have only to multiply by $10^{\text{in.}}$ whence we have, as the closest lines which can be resolved by a microscope,

$$2''.25\,\frac{1^{\text{in.}}}{n} = 0^{\text{in}}.000011\,\frac{1}{n};$$

in other words, the finest divisions which can be seen with any microscope in which the objective is "dry" are about 100,000 to the inch, while for a "homogeneous immersion" objective this number may rise to 150,000. Since the greatest value known for n in any transparent medium is about 2.5, we may say with certainty that there is no hope of ever making a greater number of lines than a quarter of a million to the inch visible.

VII. DERIVATIVE EQUATIONS

The most important of the derivative equations in the designing of optical apparatus are those which depend upon the color or variation of refrangibility of light, since they define immediately the conditions for achromatism. They are of such simplicity in our system, even when rigidly accurate, that it is doubtless worth while to give them for the sake of completeness, although it is outside the aim of this sketch of geometrical optics to develop them at length or to do anything more than indicate their use.

We will suppose the values of the indices of refraction for all the media which form the lens system to be known for a sufficient number of different wavelengths; then it will be possible to express each as a function of some one index (n) chosen as the arbitrary variable. With this as a starting-point it is easy to prove that, in general,

$$\frac{d}{dn}\mu^k \rho^k c_k = \mu^k c_k \frac{d\rho_k}{dn} + (\mu^k)^2 \rho_k \frac{dc_k}{dn};$$

whence we may write out immediately the derivative equations from the system (a) as follows: —

$$\begin{aligned}
\frac{dc_1}{dn} &= (-\gamma + \mu c')\,\frac{d\rho}{dn} + (\mu)^2 \rho\,\frac{dc'}{dn}; \\
\frac{dc_2}{dn} &= (-\gamma' + \mu' c_1)\,\frac{d\rho'}{dn} + (\mu')^2 \rho'\,\frac{dc_1}{dn} \\
&\quad \cdots\cdots\cdots\cdots\cdots\cdots \qquad (h)
\end{aligned}$$

$$\frac{d c_\lambda}{d n} = (-\gamma^{\lambda+1} + \mu^{\lambda-1} c_{\lambda-1}) \frac{d \rho^{\lambda-1}}{d n} + (\mu^{\lambda-1})^2 \rho^{\lambda-1} \frac{d c_{\lambda-1}}{d n}$$

$$\frac{d c_{\lambda+1}}{d n} = (-\gamma^{\lambda} + \mu^{\lambda} c_{\lambda}) \frac{d \rho^{\lambda}}{d n} + (\mu^{\lambda})^2 \rho^{\lambda} \frac{d c_\lambda}{d n}$$

It is obvious that the condition for ordinary achromatism is that the last equation should reduce to zero. We may leave the subject with the remark that the second derivatives, useful in the consideration of secondary chromatic aberrations, are hardly more complex than these.

APPENDIX B

THE following considerations may add something to our knowledge of this interesting phenomenon of scintillation. It is quite evident that with altitude above the earth's surface the atmosphere is neither optically homogeneous nor regularly varying in refractive power; in short, it is always more or less irregular. Imagine a point-source of light, like an enormously brilliant star, outside the atmosphere. This might be expected to illuminate the surface of the earth in a highly irregular manner on account of the infinite optical irregularity of the air. The illumination might be conceived as distributed in a mottled or reticulated manner, something like the distribution of light on a wall upon which a distant electric light is shining through a window of ordinary glass, although any valid estimate regarding the size and density of such a reticulation would seem quite hopeless. There is no such source of light; even the planet Venus at its brightest is not sufficiently bright to cast obvious shadows. But in rare instances we are able to observe the result of the illumination by a very bright linear source of light, namely, in a total solar eclipse immediately before the second and after the third contacts. It is quite evident that in this case the suppositious shadows would preserve their character only in a direction at right angles to the direction of elongation of the source. Moreover, it is equally evident that all apparent motion of these shadows would be at right angles to their lengths, since a purely linear object does not betray a displacement in the direction of its length. Consequently, such shadows, if existent, ought to appear as more or less distinct and parallel bands drifting with the layers of the atmosphere in which they have their origin, and hence with the most diverse velocities; but apparently always moving perpendicularly to their lengths. The intervals which separate them may well

have almost any value; but only a limited range of values, perhaps from a few inches to a moderate number of feet, would be likely to attract attention. This is an accurate description of the famous "shadow bands" attending total eclipses, at least of those observed by me on several occasions.

We are not limited to this particular phenomenon for proof of the general correctness of the view given in the preceding paragraph. If one could observe for a very brief time, say during a single thousandth of a second, the minute appearance of the sun, he might expect to detect variations, either in relative brightness or in sharpness of definition, in the appearance of the surface, which ought to be arranged, generally speaking, in a roughly reticulated way. But such an observation would be hopeless if restricted to the unaided eye, for, although there would be no difficulty in securing a perfectly satisfactory visual impression in the time designated, any ill-defined detail upon such a limited area as is occupied by the sun would surely escape detection. On the other hand, the use of a telescope greatly modifies the problem because then, instead of a projection of certain atmospheric irregularities from the very small area of the pupil of the eye which may be considered as practically a point, one observes that from the considerable area of the objective. Nevertheless, for those atmospheric irregularities which are remote from the observer, say from a mile or two to a hundred miles, the telescope also may be regarded as of insignificant diameter, and hence the foregoing reasoning may be extended from the case of the unaided eye to this case. In some remarkable photographs of the sun taken by Janssen a number of years ago we find observational proof of these inferences, of a most convincing character. With these photographs the exposure was sufficiently brief for the purpose, and the optical power was great enough to exhibit the fine granular structure of the photosphere. They show just such a reticulated variation in definition of details as would fit the description here given, and, as they were in no two cases alike, astronomers had no hesitation in referring the irregularities to our atmosphere rather than to a solar origin which Janssen favored.

If the general validity of the foregoing assumptions is admitted, the familiar phenomena of scintillation are easily accounted for. In the case of great and local irregularity of the

atmosphere we may be sure that even relatively small volumes of air would include variations of density, and that, owing to continuous changes, the stars would appear to the unassisted eye to undergo rapid and extreme alterations of brightness. In a large telescope the image of a star at any instant would resemble that formed by a telescope of which the material of the objective is defective; but this would change with great rapidity and thus appear to vary largely in size and shape though little in brightness. The changes would be chiefly due to bodily motion of the air, either as wind or convection currents. Chromatic changes in stellar images would be relatively insignificant, since the paths of different wavelengths of light would not be widely distributed in the neighborhood of the observer. During exceptionally quiescent states of the atmosphere these results of irregular refractions would be considerably modified. On such occasions the effective irregularities would be not only less pronounced, but also generally much further from the observer. Two consequences to be deduced from the remoteness of the source of disturbances are evident: First, the distinction between the effect as observed by the naked eye and by a telescope must largely disappear; and, second, chromatic phenomena must become more conspicuous, since an atmospheric irregularity which may greatly decrease the amount of red light that comes to the observer at a given instant lies quite remote from the path of the blue light from the same source. A third consequence is important even if less obvious. The motion of such a disturbing body of air relative to the stars will be chiefly due to the rotation of the earth rather than to the wind, because the apparent motion of the stars referred to any portion of the earth is independent of the distance of the point of reference, while the apparent motion of the wind decreases directly as its distance from the observer. The fact that such regular scintillation is observed only during moderately calm weather is a further reason for this last conclusion. The explanation of the remarkable observations of Respighi on the variations in the spectra of scintillating stars, briefly described in the text, is deducible at once from these assumptions and has already been given.

APPENDIX C

In the seventh chapter of this book the description and theory of halos has been given with considerable extension, because this class of phenomena presents us with many interesting problems, including a number still unsolved, which have engaged the attention of philosophers for centuries. The most extensive work on the subject of halos which has appeared is a memoir by Professor Bravais, contained in the *Journal de l'École Royale Polytechnique,* tome xviii., 1847. This author performs a capital service in collecting the available data of a vast number of widely scattered observations extending over a period of more than two centuries; but in the second of his aims to find a consistent physical explanation of all their features, he is far from being successful. Notwithstanding the somewhat obvious faults in the general theory here developed and applied, Bravais's work has been almost universally regarded as complete and practically final, so that for more than half a century absolutely nothing has been added or altered. We can go even further than the statement made concerning the theoretical explanations in this widely quoted work, and assert that not a single explanation originating with its author will resist critical analysis. Bravais's fundamental assumptions, which he shares in common with all his immediate predecessors in this field — namely, that elongated prisms will fall through the air in a vertical position and that flat ones will fall edgewise — are mechanically unsound. Thus only those features of halos which are independent of the ratio of length to breadth of the prisms can be possibly explained. But at the time that Bravais undertook his work everything which could be explained on the wholly correct assumptions that we have in fact, besides hexagonal crystals of ice with fortuitously directed axes, two other groups with horizontal and vertical axes, respectively, had been

completely developed. As long as the unsoundness of the additional hypotheses remained unquestioned there was no way open to account for the residual phenomena, except the assumption of the action of more complex crystalline forms. In this Bravais showed a boldness approaching temerity, for he apparently felt justified in resting an explanation upon any form not absolutely excluded by the laws of crystallography. Thus he found himself obliged to assume the presence of about a score of different forms, some being elongated prisms with intricate stellar cross-sections, others with varied pyramidal termini, and even rhombohedra. None of these has ever been observed. If any fact concerning halos can be predicated *a priori*, it is that they must be due to very simple crystals, since the number present is not only enormously great, but because any given feature would be weakened and perhaps totally suppressed by the presence of a large number of bodies suspended in the atmosphere, not engaged in its production. In the discussion of these phenomena we have assumed the existence of two types of ice crystals only, and both of these have been observed innumerable times. Moreover, they are the simplest crystal forms known for this material. In the following pages we purpose to complete and summarize the explanations given in this book.

Of all the extraordinary halos recorded the most perfectly developed and most complicated one is that described by Lowitz in the *Nova Acta Academiæ Petrop.*, tomus viii., which he observed at St. Petersburg on June 29, 1790. All three groups of directed crystals were present and effective on this occasion. Unfortunately his drawing has the defect of possessing no recognizable system of projection, and, what is still more to be regretted, there is no certain indication of the changes dependent on the sun's varying altitude, everything seen from half-past seven in the morning until noon being represented in the one diagram. Notwithstanding these shortcomings, so important to the investigator of a physical theory, they should not blind us to the fact that the observations are of remarkable merit and may well excite our astonishment that so much was seen and accurately recorded with entire lack of guidance by theory. Bravais's copy of this figure is far from satisfactory, and, as the original is not very accessible, it is quite worth while to give a faithful reproduction here, particularly since it may be

regarded as containing all the remaining features of the halo, which are certainly established by authentic observations and not included in the halo of Parry and Sabine.

The description accompanying this figure runs as follows: —

"1. The sun was surrounded by two circles *ebdk* and *eide*,

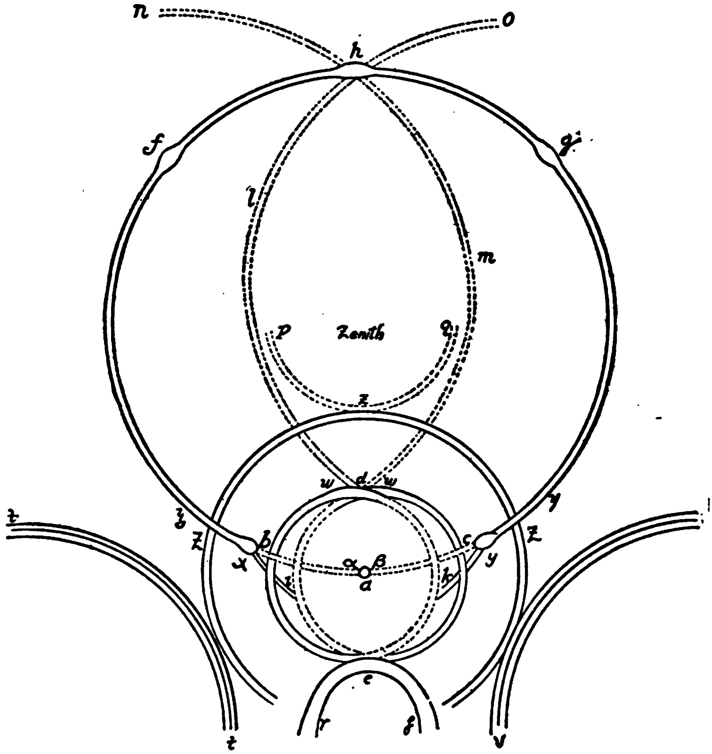

FIGURE 51.

which intersected each other at *d* and *e*, and of which the color was red on the side toward the sun and whitish on the opposite side. The sun was at *a* between their centres *α* and *β*. The exterior arcs *dbe* and *dce* were much more distinct and brilliant than the interior *dke* and *die*. The resemblance of the colors gave to these last an elongated figure or oval, and the two others had the form of a circle flattened above and below. At the upper point of their intersection a considerable part *wdw*

was observed of an extraordinary vividness, of which the splendor was almost as dazzling to the eye as that of the sun itself.

"2. The inferior point of intersection *e* touched an inverted semicircular arc *ref* very bright and wide, which was in relation to its diameter the smallest of all the arcs.

"3. Another circle *zzz*, similarly colored but at a greater distance from the sun and imperfect toward the horizon, which an arc above *pzq* gave the appearance of being horned, had the sun itself for a centre.

"4. This last circle was tangent toward the south and toward the east to two inverted arcs *tt* and *vv*, with respect to the sun, exactly like portions of a rainbow, not only from their size but also from the intensity of their prismatic colors, in which they were distinguished from all the other arcs of this meteor. Buildings prevented me from seeing whether they extended quite to the horizon.

"5. Still another circle complete and white *afhga* which enclosed a very large area, was parallel to the horizon, which it followed completely around, and which consequently had the zenith for a centre. In the circumference of this one lay the sun itself and five parhelia, of which three *f*, *h*, and *g* opposite the sun toward the northwest were white and faint, and the two others *x* and *y* at each side of the sun, in the southeast, colored and very bright.

"6. These two last parhelia which were at some distance from the intersections of the great horizontal circle by the two coronas which surrounded the sun, sent in the first place from the two sides very short colored arcs *xi* and *yk* of which the direction was inclined below the sun so as to reach the interior semicircular arcs *die* and *dke*. In the second place they had long tails, bright and white $x\zeta$ and $y\eta$, directed away from the sun and included in the circumference of the great circle *afhg*.

"7. Finally two great circular arcs *dlh* and *dmh* of a white color appeared inside the great horizontal circle *afhg*, but they were so faint, that several persons to whom I attempted to show them, were unable to perceive them. In the first place they met near the sun at *d* in the dazzling brightness, and they again crossed each other on the other side, as well as the great circle, in the centre of the pale parhelion *h*, whence they extended

sensibly further, several degrees beyond the great circle toward the northwest point of the horizon, to *n* and *o*.

"Such was this beautiful meteor at ten o'clock in the morning when it had attained its greatest perfection. With regard to the successive changes of its parts, I made the following observations in addition:

"At half-past seven in the morning the two coronas around the sun *dbek* and *dcei* were not yet perfect, only the interior arcs *die* and *dke* appeared under the form of a perfect oval with a bright region above at *d*. It was only little by little that this brightness extended itself on both sides, in the form of the arcs *dw*, *dw*, which became larger and larger, until they united finally at nine o'clock at *e* near the brilliant semicircle *ref*.

"That which was most remarkable in these two circles or coronas, which intersected each other, was the fact that after having attained their perfection they approached each other more and more until in the end they formed only a single corona which had the sun at its centre. Nevertheless the upper and lower portions always retained a very sensible brightness. About this time the two arcs with prismatic colors *tt* and *vv* disappeared. The two parhelia *x* and *y* on the contrary separated themselves more and more widely from the sun and finally disappeared entirely at a quarter to eleven.

"The great horizontal circle *afhga* and its three faint parhelia *f*, *h*, and *g* still remained and did not disappear until eleven thirty-five. That which still deserved being remarked in this great circle, is that preserving always the sun and its five parhelia in its circumference, it remained constantly parallel to the horizon, and consequently retained the zenith as a centre; this is why the more the sun approached the meridian, the more it rose above the horizon, the great circle of which the extent was at first extremely large, when the sun was more elevated, diminished continuously, until at the end its diameter became almost as small as that of the two united coronas. The same thing happened to the two arcs *dlh* and *dmh* which were within the horizontal circle. Their points of intersection were always at *h* and *d*.

"At noon there remained nothing of this beautiful phenomenon except a simple corona very bright at its upper and lower portions, which finally disappeared entirely at half-past twelve.

"In general this meteor was composed of twelve arcs, of which nine were colored, in such a way that in all cases the edge toward the sun was red and the opposite edge white."

There is comparatively little new in this remarkable halo, that is, little that has not been already described and explained. Section 1 is somewhat obscure in ascribing the form of the inner curve to the resemblance of the colors, but it becomes quite clear if we interpret it as meaning that the two arcs named seemed from their similarity to belong together. This was indeed the case because they were halves of the ordinary twenty-two-degree circle. There is nothing surprising in the fact that this appeared to the observer to be distinctly oval, for in a great number of similar cases the same error is found in the records. Even the present writer, although quite aware of the deception, found a well-developed circumscribing oval, seen at New Haven during the early afternoon of April 13, 1901, sensibly round, with the enclosed twenty-two-degree circle — here also the less brilliant of the two — appearing as an ellipse. In passing, it may be noted that this observation suggests the possibility that many of the high-sun halos, in which the upper and lower portions are decidedly brighter than the rest, are really the circumscribing oval unaccompanied by the ordinary twenty-two-degree circle. This is a point which may be worthy of attention in the future.

Section 2 describes a rare feature, although recorded by several other observers. It is, without doubt, the inferior complement of the arc *xev* of Parry and Sabine, the explanation of which we have found in the action of the crystals of the *B* group. Its most characteristic peculiarity was the smallness of its apparent diameter; but its brightness pretty certainly indicates fixity of direction of its faces. The record is hardly precise enough to admit of quite satisfactory proof by calculation, nevertheless it is worth stating that an approximate calculation for an altitude of the sun of 30°, corresponding to about the beginning of Lowitz's observations, gave me a radius of about 6° for the summit of an arc more nearly resembling a parabola than a circle. The angular distance of the apex from the sun was found to be 24°.7. When we take into account the inevitable excess in eye estimates of angular magnitudes

of objects near the horizon, this is not at all a bad agreement with the drawing, although the radius of the curve is there approximately twice this value. A very moderate oscillation of these crystals in their descent would materially increase the radius of the arc.

Section 3 describes the well-known forty-six-degree circle with its upper tangent arc. In this halo the evident paucity of undirected crystals, as evinced by the faintness of the twenty-two-degree halo, suggests a question concerning the origin of the greater circle. It is true that fortuitous crystals should produce such a circle, but whether with sufficient intensity to make it visible against a bright sky is quite another question. I have never found a trace of the forty-six-degree circle, although I have observed scores of simple halos, not a few of which were remarkably bright. On the whole, in view of the facts (*a*) that there are six points in the circumference at which one or another of the directed crystals may contribute largely to the light from it, and (*b*) that a slight rocking motion of the crystals at these points would produce extended arcs coincident with the circle, I am inclined to replace the accepted explanation by referring the phenomenon to concurrent action of all three groups of directed crystals, eliminating the fortuitous crystals as wholly negligible agents.

Section 4 gives us nothing new, since these arcs are the same as those seen by Parry and Sabine at a slightly lower altitude of the sun. They appear to have vanished about nine o'clock, at which time the sun stood 40° above the horizon.

Section 5 is a description of the parhelic circle of which the explanation has been given quite fully in the text.

Section 6 presents nothing new except the singular oblique arcs extending from the parhelia to the twenty-two-degree circle. No one else has described them with the same particularity, hence they are very properly known as the "oblique arcs of Lowitz." Galle has explained them satisfactorily by referring them to the action of crystals similar to those producing the neighboring parhelia, but possessing a rocking or balancing motion in their fall.

Section 7 describes the singular oblique arcs which have proved such a puzzle to all writers on this subject. According to our views they are the visible portion of a continuous curve

which owes its origin to two refractions separated by an internal reflection by crystals of class *B*. A special solution for the case of the sun at an altitude above the horizon of 65° has been shown in Figure 39, page 152. The values derived by calculation for this particular halo are given in the table immediately following: —

A 60°	90°	110°	130°	133°	140°	150°	160°
S' 27.6 31.4	46.1 27.7	62.0 24.1	83.3 19.8	86.2 18.9	95.4 17.2	110.5 14.7	129.8 12 6
S'' 92.4 31.4	136.1 27.7	158.0 24.1	176.7 19.8	179.8 18.9	184.6 17.2	189.5 14.7	190.2 12.6
S''' 92.4 43.1	136.1 37.6	158.0 32.3	176.7 26.4	179.8 25.1	184.6 22.8	189.5 19.4	190.2 16.7

Here the figures in the first horizontal row are the differences between the azmuth of the sun and the vertical plane containing the poles of the refracting surfaces; consequently, also the angle between the vertical containing the sun and the vertical base of the prism upon which the internal reflection takes place. The rows attaching to S' give the azmuth and zenith distance of the image of the sun formed by the first refraction, the upper figure being the azmuth angle. The next two rows give the co-ordinates of the images formed by the subsequent reflection, and the rows attaching to S''' are the corresponding co-ordinates for the images formed by the completed action of the prisms, and are the points through which the double spiral of Figure 39 is drawn.

It is very difficult to apply the theory with any degree of confidence to these observations of Lowitz because of their indefiniteness. For example, the observer remarks that his description applies to the state at ten o'clock in the morning, at which hour the zenith distance of the sun was only 43°. This would have brought the zenith well within the outer concentric halo, which ill accords with the drawing. At nine o'clock the sun's zenith distance was 50°, which still seems too high for the drawing. It is evident that we have a wide range of choice for the elements of our problem, with little hope of reproducing the composite drawing in the calculations founded upon these elements;

but if we adopt 45° as the zenith distance of the sun, the result of calculations founded on the theory presented in this book is as follows: The complete curve consists of two spirals like those of Figure 39, which together form two loops, one including the other, with the node at the anthelion. In the present case, however, the inner loop is far larger, embracing the zenith and tangent to the twenty-two-degree circle at its highest point. The point of tangency is not a cusp, as represented in the drawing, but the lowest point of a nearly circular arc of about twenty-nine-degree radius — a difference which counts for nothing, since this portion of the curve would be quite invisible on account of the overwhelming brightness of the arcs *wdw*. The two branches of the curve intersect at an angle of very nearly 60°, which agrees with the drawing. From the anthelion the two branches extend with continuously diminishing brightness and curvature until they meet at a point 34° below the sun. The inner loop of the calculated curve embraces somewhat more area with respect to that surrounded by the parhelic circle, but this is chiefly due to the difference in shape in the region between the zenith and the sun. It is a singular result of the calculations that, whereas the spirals of Figure 39 were produced by light incident on one of the upper oblique faces of the *B* crystals, in this case only the outer loop is thus formed, the inner one being produced by an entering refraction on the upper horizontal surface and a final emergence from a lower oblique face.

As this completes the explanation of all known features of the complex phenomenon called the halo, it may be well to collect them in tabular form. We will first give those of which the origin has been known for a longer or shorter time, with the name of the physicist who first found the true explanation.

1. Halo of twenty-two-degree radius. MARIOTTE.
2. Parhelia of 22°. MARIOTTE.
3. Oblique arcs of Lowitz. GALLE.
4. Tangent arcs to twenty-two-degree halo, which become the circumscribing oval with high sun. YOUNG and VENTURI.
5. Halo of forty-six-degree radius. CAVENDISH. (Unless objections given on page 219 in regard to this feature are valid.)
6. Horizontal tangent arcs to forty-six-degree halo. GALLE; perfected by BRAVAIS.

To these must be added the following which have not hitherto been explained at all, or wrongly explained because grounded upon theories which are untenable : —

7. Lateral tangent arcs to the forty-six-degree halo.
8. Parhelic circle.
9. Paranthelia.
10. Anthelion.
11. The arcs above and below the twenty-two-degree halo.
12. The short oblique arcs through the anthelion.
13. Spiral arcs through the anthelion.
14. Vertical columns.

There is, however, a celebrated halo that contains a feature not mentioned in the list, which has given a great deal of trouble to writers on this subject from the time of Huyghens down. It is a rather remarkable halo observed by Hevelius in 1661, and described fully in Smith's Opticks, vol. i., pp. 221, 222, although with the exception of this feature it seems to have been a well-developed halo depending upon the presence of the *A* group for its chief characteristics. The exceptional feature is a circle, of which only the lower portions are shown in the figure illustrating it, everywhere 90° from the sun, and therefore a great circle. Bravais, who styles this as the most authentic of all extraordinary halos, cites all the explanations offered, points out their fallacies, but quite frankly declares his inability to propose any more satisfactory theory. Since I am forced to follow Bravais exactly in this respect, it may be well to review the evidence of the existence of the ninety-degree circle, beyond that contained in the original record. There is nothing in the records of the time since Bravais which bears upon this point, at least a search by me has led to no result; hence we are confined to the three examples which that author finds.

The first is found in the description of the halo observed at Melville Island, given by Parry and Sabine. The passage in the last paragraph of the quotation, page 143, describing the faint light about a quadrant from the sun, is taken as an observation of the circle in question; but a most casual reading demonstrates that such an interpretation is an entire misapprehension.

The second instance is found in a very uncritical description of a halo seen at Derby in England, in 1802, and published in the "Philosophical Magazine," vol. xii., p. 373. In this case neither the name of the observer nor the place of the sun in the heavens is given. The passage in which Bravais finds evidence of the ninety-degree circle reads as follows: ". . . the fourth [circle] circumscribed all the others, and was touched upon the western side by part of another of the same diameter." It is quite clear that this circle did not have a radius of 90°, not only because no ordinary observer would dream of calling a great circle of which the sun occupies the position of one pole, a circumscribing circle, but also because in that case another circle tangent to it and of the same diameter would be identical with it. Unquestionably, this fourth circle was the forty-six-degree halo, and the circle touching it was the upper tangent arc.

The final case appears to be much more conclusive. It is that of a lunar halo observed by Erman in Siberia, in 1828.[1] Here, with the most minute particularity, that traveller gives the results of his observations, together with the fact that at 10.30 P. M., Tobolsk mean time, the measured distance of the moon from the vertex of an auroral arch was 83°.2; moreover, that at the same instant the lunar halo intersected the auroral arch a few degrees to the west of its vertex. This seems very convincing as to the existence of a halo with a radius of 85° to 90°; but reference to the details of the original account shows certain peculiarities which cannot fail to awaken strong doubts concerning this conclusion. In the first place, Erman describes the halo without any intimation that it is an unusual one. Then he mentions the fact that it coincided with a part of one of a system of concentric arcs which are supposed to be auroral on account of their fixity of position with respect to the earth. Finally, he gives the measured distance of the moon from the apex of the lowermost arch at 6.30 in the evening, which he found to be 86°. At this time the moon was close to the horizon; consequently, if the radius of the halo was 90°, it would have intersected all the auroral arches nearly orthagonally, and a partial coincidence at any point would have been quite out of the question. But this is not the only incon-

[1] Erman, *Reise um die Erde*, vol. i., p. 544.

sistency. An investigation as to the position of the moon at the place given and at the epoch of November 24, 1828, 10.30 P. M., shows that its true distance from the point indicated as marking the place of the vertex of the auroral arch was 107°; hence Erman's statement is erroneous.

But it is quite easy to supply to the printed account an emendation which eliminates all the difficulties and contradictions. We find that, on the evening in question, the distinguished traveller was at Sawodinsk, a place 2° north of Tobolsk, engaged in making a complete and protracted set of observations on the magnetic elements of the place. During the intervals of these observations — important as a part of a very elaborate system — he entered in his notebook the contemporary phenomena of auroral arches and the halo. At the later hour named he made the angular measure, probably with a sextant. So much is certain. Now let us suppose that he chose the easy task of measuring the distance between the summit of the auroral arch and the nearest point of the halo instead of the less simple task of measuring the interval separating this summit from the relatively brilliant moon, in which case he would have been obliged to experiment with the dark glasses which are not well adapted for this kind of work. Under this supposition and the assumption that the halo was the ordinary one of 22°, we find that the distance separating the apex of the arch and the moon was 105°.2, which accords well enough with the astronomical fact. The only other modification necessary is to assume that the circle which intersected the auroral arch a few degrees to the west of its vertex was the vertical circle through the moon instead of the circle which accompanied the moon. With these highly plausible assumptions the records of a trained observer are made perfectly clear and probable, while without them they are entirely self-contradictory; yet with these modifications the last bit of confirmatory evidence for the ninety-degree halo of Hevelius falls to the ground. It does not seem unphilosophical to conclude that an inexplicable phenomenon recorded only once in a quarter of a millennium does not really exist.

THIS BOOK IS DUE ON THE LAST DATE
STAMPED BELOW

AN INITIAL FINE OF 25 CENTS

WILL BE ASSESSED FOR FAILURE TO RETURN THIS BOOK ON THE DATE DUE. THE PENALTY WILL INCREASE TO 50 CENTS ON THE FOURTH DAY AND TO $1.00 ON THE SEVENTH DAY OVERDUE.

MAY 18 1945	
JUN 7 1946	
2May'49	
'50	

LD 21-100*m*-12,'43 (8796s)